VISUALIZATION OF CAREER-RELATED COMPUTER-MEDIATED COMMUNICATION FOR INCREASED KNOWLEDGE MANAGEMENT

THESIS

Marc A. Grayson, 2nd Lieutenant, USAF

AFIT/GIR/ENV/03-05

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

AFIT/GIR/ENV/03-05

VISUALIZATION OF CAREER-RELATED COMPUTER-MEDIATED COMMUNICATION FOR INCREASED KNOWLEDGE MANAGEMENT

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Information Resource Management

Marc A. Grayson, B.S.

Second Lieutenant, USAF

March 2003

AFIT/GIR/ENV/03-05

VISUALIZATION OF CAREER-RELATED COMPUTER-MEDIATED COMMUNICATION FOR INCREASED KNOWLEDGE MANAGEMENT

Marc A. Grayson, B.S.
Second Lieutenant, USAF

Approved:

_____// Signed //_____	____10 Mar 03__
Maj. Mark Ward, PhD (Chairman)	date
Associate Professor, Information Resource Management	
Department of Systems and Engineering Management	
_____// Signed //_____	____10 Mar 03__
Lt. Col. (sel) Karl Mathias, PhD (Member)	date
Associate Professor, Computer Science and Engineering	
Department of Systems and Engineering Management	
_____// Signed //_____	____10 Mar 03__
Dr. Alan Heminger (Member)	date
Associate Professor, Information Resource Management	
Department of Systems and Engineering Management	

Acknowledgements

First and first most, I acknowledge the assistance of all committee members whom each have played a major contribution to the completion of this research. Maj. Ward gave me the opportunity to be flexible and creative in the direction of this thesis. He was not only an advisor, but a mentor as well. Dr. Heminger provided the fundamentals of knowledge management and the reasons why this research must include KM support. Lt.Col. Mathias was a mentor and a motivator through my challenges with database track of electives. These classes were instrumental to the successful completion of the OrgDiscovery, designed for this study. Dr. Chen of the University of Arizona was a mentor and provided much guidance in my academic and career endeavors. My experiences with him at the Artificial Intelligence (AI) Lab provided ideas and technologies that were applicable to this study. I'd like to thank Lt.Col. Biros for his faith in my research and assistance in completing a conference paper.

I would like to acknowledge my family. My parents have provided me with many life opportunities due to their unconditional love and support. I'd like to thank my brother, for being a role model as an officer in the United States Navy. I'd like to thank my fiancé, for being supportive through this whole process, through all my high's and lows. Finally, I'd like to thank God for without him, such accomplishments are not possible.

Table of Contents

Page

Page

List of Figures

List of Tables

Page

Abstract

Retention and retrieval of organizational memory has been the concentration of many conceptualized models of an organizational memory information system (OMIS). This thesis presents an extended view for system development of an OMIS from a knowledge management perspective. The United States Air Force maintains various career-related mailing lists (listservs) for information technology (IT) specialists sponsored by the Air Force Communications Agency (AFCA). AFCA has realized the importance of monitoring the communication for patterns in content and behavior. This thesis details an experimental study, which includes a repository of computer-mediated communication (CMC) of IT specialists, analyzed by software created for this study, the OrgDiscovery system. This system is designed to visualize the content and behavior patterns of computer-mediated communication. The purpose of this study is to show that visualization of mailing list communication provides a more usable method to make conclusions about the participants of mailing lists versus the text-based Microsoft Outlook. M.S. Outlook is the mail program currently being used by management to store and review mailing list emails.

VISUALIZATION OF CAREER-RELATED COMPUTER-MEDIATED COMMUNICATION FOR INCREASED KNOWLEDGE MANAGEMENT

I. Introduction

Background

There has been an explosion of organizational communication through the advances of computer-mediated communication (CMC). Companies, organizations, and work groups are utilizing communal forums to transfer tacit knowledge, job know-how and experiences, to achieve objectives (Gore & Gore, 1999). There is an increased business reliance for "…computer-supported forms of communication, collaboration, and coordination" (Trauth & Jessup, 2000, p44). Davenport and Prusak state, "…research shows that knowledge is most effectively communicated through a convincing narrative that is delivered with formal elegance and passion" (1998, p.81).

Computer-Mediated Communication (CMC) is a forum where members can communicate freely with peers of a particular community (Ahuja, 1998). The informality allows for the free flow of knowledge to peers that may not be found through face-to-face communication. Tacit knowledge, individual's know-how, mental models, beliefs, and perceptions (Gore & Gore, 1999) is knowledge that comes with experience. Explicit knowledge refers to the "…codifiable component that can be disembodied and transmitted" (Hahn & Subramani, p.303, 2000). CMC establishes an environment where tacit knowledge can be converted and codified into explicit knowledge (Nonaka & Konno, 1998). Computer-Mediated Communication (CMC) can be used as a medium

where individuals in various locations can communicate job and career-related issues (Nonaka & Konno, 1998).

It is important for managers who establish career training, documentation, tools, and techniques for those in a given mailing list community, to be aware of the issues of those participants. An archive of computer-mediated communication "…contains rich information about both the content and the behavior of participants" (Zhu &, Chen, 2001, p1). Many CMC systems have focused on the organization of discussion content through the mediation in which participants communicate (Nunamaker, et al, 1991). The Answer Garden, a content-based system, makes codified knowledge retrievable and those with such knowledge accessible (Ackerman, 1998). Many visualization techniques have been developed to summarize behavior patterns of CMC, e.g. Loom (Donath et al., 1999), Chat Circles (Donath et al., 1999), and PeopleGarden (Xiong & Donath, 1999). Very few systems like the Communication-Garden system (Zhu & Chen) have been developed to analyze and visualize both the content and behavior patterns of this type of communication.

Implications

The Communication-Garden System (Zhu & Chen, 2001) was shown successful in providing users assistance in understanding both the behavioral patterns and content of the participants during a CMC process. This study further validates such research. The OrgDiscovery system, the prototype system developed for this study, attempts to help users understand the content as well as the behavioral patterns of mailing list communication of specialists. The OrgDiscovery system acquires such capability

through rigorous preprocessing, analysis, and visualization of CMC. Finally, usability testing, comparing user evaluation of OrgDiscovery against the current PC email program, Microsoft Outlook, provided a platform to evaluate if the OrgDiscovery tool is more effective in achieving the organizational knowledge that management desires in comparison to the current system being used.

Theory for Knowledge Management Support Approach

There are many approaches to consider why and how to manage knowledge. For this study, it is important to understand the basic considerations in managing the knowledge found inclusive in the mailing list discussions. Hahn and Subramani state that the two basic considerations in managing knowledge are "(1) where the knowledge resides and (2) the extent to which the knowledge is structured" (2003, p304). The dimensions categorize the various knowledge management systems used for specific knowledge management support.

The fourth cell as shown in Figure 1, as related to this study, categorizes electronic communication as a free, unstructured platform to communicate and transfer knowledge. As noted by Hahn and Subramani, instances of such systems provide a platform for questions to be posted as well as employee response with answers or suggestions (2000). This framework notes that such communication is unstructured and is dependant on the participation of individuals. Since such communication is unstructured, but is rich in organizational memory (Hahn & Subramani, 2000), this study hopes to highlight a need for developing a more **"structured"** environment for analysis

and visualization of both the content and behavioral patterns of computer-mediated communication.

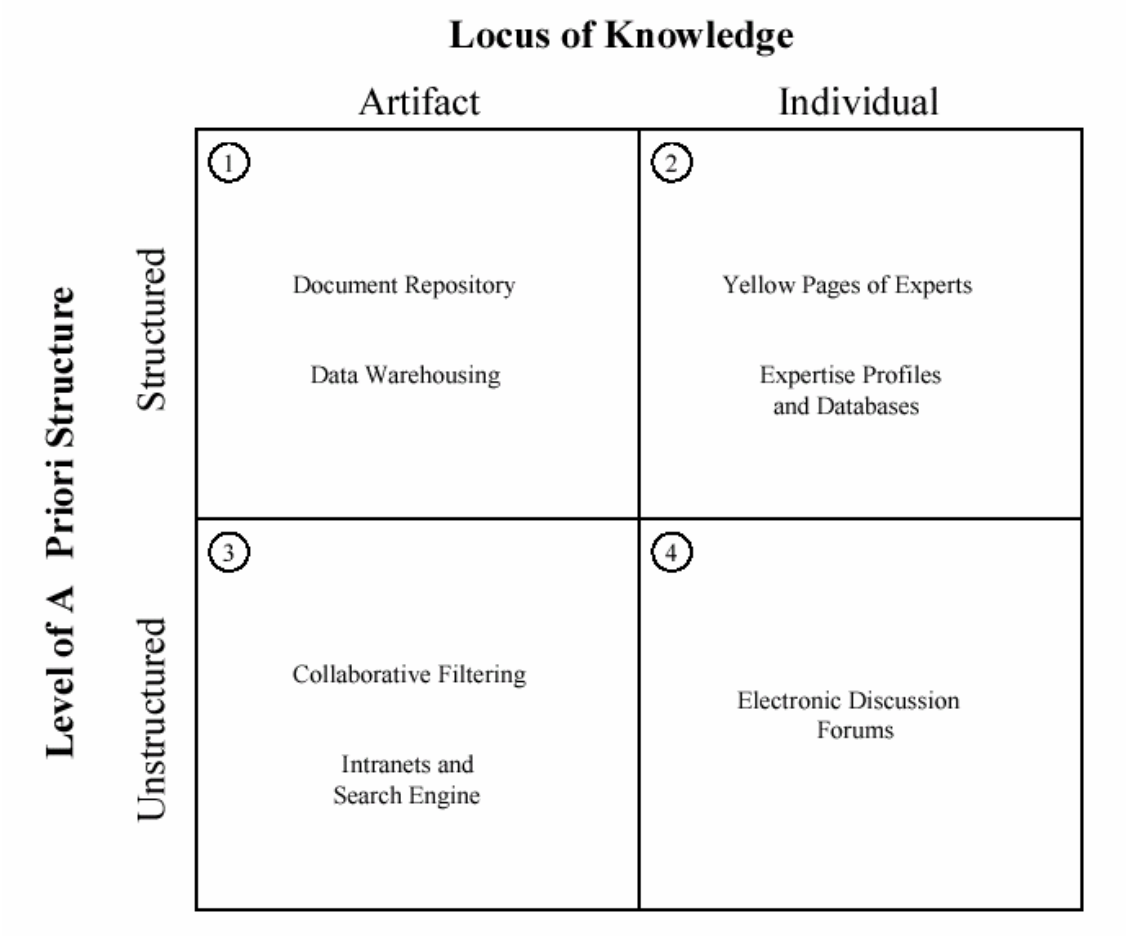


Figure 1: Framework for Knowledge Support

Note. Modified From "A Frame Work of Knowledge Management Systems: Issues and Challenges for Theory and Practice" (p. 304), by J. Hahn and M.R. Subramani, Copyright © 2000, Atlanta: Association for Information Systems.

Problem Statement

Much work has been done to provide new ideas to capture, organize, and visualize the content and behavior patterns of computer-mediated communication (CMC). This importance has been shown, but little has been shown to effectively collect and retain this resource for reusability and analysis of collective organizational knowledge. Research has provided information analysis techniques for the indexing, clustering and visualization for search and retrieval of collaborative content of group support systems (GSS) (Chen, et al, 1998). This study aims to look at one area where organizational knowledge is shared, mailing lists (listservs), and demonstrate the methods used in capturing, analyzing, and visualizing the knowledge as well as behavioral patterns of the participants. The techniques shown useful in the field of information retrieval and visualization, may be used to not only retain and retrieve electronic communication, but could provide a tool for conducting a study of an organization's behavioral patterns and interaction in given community.

Research Questions

Can an organizational memory information system (OMIS) be built to provide an understanding of behavioral and content patterns found in mailing list communication through visualization? Can this system be shown as more usable by end-users than the current software, Microsoft Outlook, used to identify such patterns?

Significance of this Study

Retention and retrieval of organizational memory has been the concentration of many conceptualized models for organizational memory information systems (OMIS)

(Hackbarth & Grover, 2000). This research presents an extended view for system development of an organizational memory information system from a knowledge management perspective. Computer-mediated communication (CMC) has been viewed as rich in organization memory (Hahn & Subramani, 2000), and as a method for transferring organizational memory in large volumes (Grayson et al, 2002). Managers of employees reviewing such communication need a system to provide clues to further support organizational effectiveness. The techniques examined to obtain organizational knowledge from an archive of mailing list emails include: analysis and pre-processing of organizational data, text mining of narratives through Natural Language Processing (NLP) techniques, and evaluation/interpretation of processed content through summaries and visualization. This study includes a repository of computer-mediated communication (CMC), analyzed by the prototype system, OrgDiscovery. This system will be turned over to Air Force Communications Agency (AFCA), to conduct future research of career-related mailing lists discussions.

Scope and Assumptions

The Air Force Communications Agency (AFCA) has established a "Comm-Info Arena", where people can join from among the 33 Comm-Info career-related mailing lists. Each mailing list represents a distinct career-field where sharing of knowledge is performed by shared work practitioners, "people doing similar work but in different settings" (Markus, 2001). Some of the mailing lists include DIALTONE, telephone systems, METENAVS, metenav systems, WIREDAWG, cable and antenna systems, and WM, workgroup managers, to name a few. The number of people on each list range from

50 to over 300 people. AFCA is not only the moderator of these lists, but is also involved with the training, documentation, and techniques of the specialists on the mailing lists. Thus AFCA has an invested interest in the conversations that take place publicly on each mailing list, content as well as the behavioral patterns of the participants. The questions and answers being discussed on CMC could have a correlation to training, techniques, tools, and resources provided by the overseeing organization, (Grayson et al, 2002).

This study includes a repository of the AFCA career-related mailing lists' discussions, where the narratives include questions and answers for job-specific tasks. Mailing list members can provide multiple responses (solutions) to any one particular question. Members can refrain from responding when all possible solutions have been expressed. AFCA has noted that there is no existing repository of the mailing list discussions and no current on-line resource to conduct a search of an archive of mailing list responses. Thus, I joined all the mailing lists (minus the restricted lists) as a member on December 6, 2001. The communication from these lists has been stored in an email account, where each email is stored into a mail folder according to the mailing list the message came from. This collection was performed for approximately 1 year.

Thesis Overview

The next chapter examines the constructs of knowledge, knowledge management, organizational memory, organizational learning, scenarios of knowledge reuse, computer-mediated communication (CMC), text mining via natural language processing (NLP), and previous research performed to study and analyze a repository of computer-mediated communication. Chapter three will provide the methodology taken to produce an

organizational memory information system (OMIS) using a repository of shared CMC as input, and how to measure its usability for practitioners. Chapter four will present the results of usability testing that compares OrgDiscovery to Microsoft Outlook. The results of the usability testing between the OrgDiscovery system and the current PC mail program used by managers and specialists to store the mailing list emails will be presented. Finally, chapter five discusses the conclusions drawn from the research, limitations of the current study, and directions for future research in this area.

II. Literature Review

Introduction

This chapter examines previous research that has guided the process in building an organizational memory information system (OMIS), developed to organize content and visualize behavioral patterns within computer-mediated communication (CMC). The literature review includes reasons for building such a system: knowledge, knowledge management, organizational memory, and organizational learning. This chapter also reviews the notion of an organizational memory information system (OMIS) and past research that provides conceptual and actualized systems. Previous research is presented on computer-mediated communication (CMC), the input utilized by the OMIS built for this study. Research on conversation visualization, natural language processing (NLP), and knowledge discovery in databases (KDD) is reviewed specifically for organizing content and visualization of behavior patterns within a CMC process. Finally, an overview of research on usability and technology acceptance (TAM) of software for end-users concludes the review of prior research applicable to this study.

Knowledge

Knowledge, in general terms, can be viewed as a "…fluid mix of framed experience, values, contextual information and expert insight that provide a framework for evaluation and incorporating new experiences and information" (Davenport & Prusak, 1997, p5). Knowledge has been shown to be an intangible asset to organizations effectiveness. Nonaka states that making individual knowledge accessible to others helps

organizations create a knowledge-creating organization (1998). There are two kinds of knowledge: explicit knowledge and tacit knowledge. Explicit knowledge "…can be expressed in words and numbers and shared in the form of data, scientific formulae, specifications, manuals and the like" (Nonaka & Konno, 1998, p42). Explicit knowledge can be also be expressed and transferred through computer-mediated communication (CMC) (Hahn & Subramani, 2000). This type of knowledge can be easily transferred between individuals through various mediums. Explicit knowledge is "…that which is already documented; located in files, manuals, databases, etc" (DeTienne & Jackson, 2001, p3).

Tacit knowledge is "…highly personal and hard to formalize, making it difficult to share with others" (Nonaka & Konno, 1998, p42). Hahn and Subramani view tacit knowledge as the "…rooted know-how that emerges from action in a particular context" (2000, p303). It consists of "…mental models, beliefs, and perspectives so in ingrained that we can take them for granted" (Nonaka, 1998). Tacit knowledge, as expressed by Lubit, must be widely transferable within an organization (2001). It also consists partly of technical skills – "…the kind of informal, hard-to-pin-down skills captured in the term 'know how'" (Nonaka, 1998, p28).

Organizational knowledge is created through cycles of combination, internationalization, socialization, and externalization that transform knowledge between tacit and explicit forms (Nonaka, 1994). This type of knowledge is seen as a key resource that should be captured and dispersed across the organization to create an organizational advantage (Hahn & Subramani, 2000, p302). Gore & Gore view organizational knowledge as comprised of "…corporate knowledge and shared understandings, but it

has similar characteristics to individual knowledge" (1999, p554). Organizational knowledge is created by means of information and social interaction for furthered development (Gore & Gore, 1999).

Nonaka refers to four basic patterns for creating knowledge in any organization. All four patterns exist dynamically, in no particular sequence (Nonaka, 1998).

1) Tacit to Tacit – This is the interaction between one individual sharing knowledge directly with another. Socialization is a limited form of knowledge creation. If it does not become explicit, it cannot be easily leveraged.
2) Explicit to Explicit – An individual can combine discrete pieces of explicit knowledge into a new entity. This combination, however, does not extend the organizations existing assets, rather brings together existing information from various sources.
3) Tacit to Explicit – Includes the ability to articulate the foundations of tacit knowledge into explicit to be shared with members of an organization. Another perspective is the innovative approach to situations based on the tacit knowledge developed over the years in the job.
4) Explicit to Tacit – As new explicit knowledge is shared and internalized by members of the organization, they can extend and broaden their own tacit knowledge base (Nonaka, 1998).

Hahn & Subramani refer to computer-mediated communication as a forum where tacit knowledge becomes explicit. Once explicit, the knowledge embedded within communication is readily transferable. To fully leverage this knowledge is the need for management.

Knowledge Management

Knowledge management is the fundamental concern with achievement of knowledge moving from the domain of the individual to the organization for organizational progress (Gore & Gore, 1999). Knowledge management includes making knowledge easily accessible; this is accomplished through knowledge being presented to a receiver in an explicit manner. Managing knowledge is a complex process. Many knowledge management efforts are limited to creating electronic means to foster knowledge transfer and knowledge storage (Lubit, 2001). Chait describes an effective knowledge management system as including the management of four domains: content, culture, process, and infrastructure (2000).

1) Content – There must be an understanding of what content is used, by whom, and the relative importance of each element to individuals, groups, and organization objectives.
2) Culture – There must be a business impact to make knowledge management a worthwhile and time-worthy endeavor.
3) Process – Identification of how to manage the knowledge – how to capture, evaluate, pre-process (cleanse), store, provide, and use the knowledge.
4) Infrastructure – There must be teaching, training, coaching and support to ensure knowledge management applications are used effectively. (Chait, 2000)

Chait provides the understanding that knowledge management endeavors must have purpose. Knowledge management as demonstrated above, must have a clear understanding of the audience who requires given knowledge. The knowledge management process is dependent on how the knowledge is managed, along with an

infrastructure to ensure it's effective. Knowledge management activities depend on users utilizing its functionality to its greatest extent. Knowledge management systems should provide the means to spread best practices, and bring together pieces that cause innovation to new services, techniques, and products (Lubit, 2000). Knowledge, as explained by Lubit, has the ability to spark new innovative practices by leveraging on what was done before (Lubit, 2000). This is critical when creating a system that can capture and present past knowledge for personnel to make better and more knowledgeable decisions.

Organizational Memory

In its most basic sense, Walsh and Ungson refer to organizational memory as "…stored information from an organization's history that can be brought to bear on present decisions" (1991, p.61). Hackbarth states a repository of organizational memory is provided by "…individual roles and relationships of these roles" (Hackbarth & Grover, 1999, p24). The memory of individuals can be developed and exploited through shared functions and activities caused by these relationships. Stein and Zwass note, "…organizational memory is an instance of collective memory" (2001, p88). This denotes that memory is utilized at a given moment by drawing on experiences from the past. Experience and participation in various organizational functions develops a history that is particular to an individual or a group as an entity. "People who are knowledgeable not only have information, but also have the ability to integrate and frame the information at hand within the context of their experience, expertise, and judgment" (Hackbarth &

Grover, 1999, p21). Organizational memory is means by which knowledge from the past influences present organizational activities (Stein & Zwass, 2001)

Walsh and Ungson provide three basic assumptions about organizational memory that has guided the development of organizational memory information systems (OMIS) and various knowledge management systems. They note that (1) organizations functionally resemble information-processing systems that process information from the environment. Since people resemble information-processing systems, systems should have the intelligence to mimic the mindset of people. (2) They also depict organizations as interpretative systems. Organizations, like interpretive systems, have the ability to interpret and make judgments based on the information provided. (3) "Memory is a concept that an observer invokes to explain part of a system or behavior that is not easily observed…this is the basis that underlies the concept of interpretation systems" (Walsh & Ungson, 1991, p60). Individuals can explain and express thoughts or actions that may not be observed, but are invoked from past experiences.

Computer-mediated communication (CMC) has been viewed as rich in organization memory (Hahn & Subramani, 2000). If organizational memory can be brought forth to help make present decisions (Walsh & Ungson, 1991), then computer-mediated communication (CMC) should not only be stored and captured, but reused to make decisions to ad hoc problems.

Organizational Learning

Levitt and March view organizations as learning by "…encoding inferences from history into routines that guide behavior" (1988, p319). Research on organizational

learning has included work towards increasing the transfer of knowledge across organizations (Argote, 1999). Knowledge is best leveraged for organizational effectiveness when it is transferred for organizational learning. Argote explains that current research examines whether organizations learn from experience of other organizations, or if one organization benefits from knowledge acquired at another (1999). "Organizations capture the experience of other organizations through the transfer of encoded experiences in the form of technologies, codes, procedures, or similar routines," (Levitt & March, 1988, p329). As with computer mediated communication (CMC), organization memory is transferred when it is coded and made explicit.

Also, distributed units are likely to have similar solutions and problems (Goodman & Darr, 1998). Advances in technology and information systems have increased the potential for facilitating knowledge transfer and organizational learning across geographically distributed sites (Goodman & Darr, 1998). Furthermore, computer-mediated communication (CMC) has allowed for knowledge to be transferred beyond the walls of distinct and homogeneous organizations (Hahn & Subramani, 2000). CMC supports the use of peers offering suggestions to job-related problems or issues, followed by peer review of those suggestions in subsequent responses (Hahn & Subramani, 2000, Grayson et al, 2002). Organizational learning as found in a CMC process is due to the interaction found amongst participants of this type of communication. This interaction of peer-review found in a CMC process is a form of lower-level learning (Stein & Zwass, 1995), where peers continuously review responses until the best solution or method is found (Grayson et al, 2002). Lower-level learning as described by Stein and Zwass "…best takes place if the standards are part of the memory

of the organization so that its members can detect and correct variances" (2001, p86). Since multiple members of CMC can respond to questions and subsequent responses alike, members can correct variances in responses to enhance lower-level learning for all members in a particular community.

As seen with the communities utilizing CMC, the dynamic process of knowledge creation and organizational learning links individuals and groups with similar tasks. These links are seen as "communities of practice", where they play an important role in communicating knowledge and provide an environment for organizational learning (Brown & Duguid, 1991). These communities of work group practitioners, people with same job specifications but reside in different settings (Markus, 2001), have their "…own unique language and context-specific vocabularies" (Hahn & Subramani, 2000, p303). This plays an important factor on how an organization memory information system is developed to facilitate the organizational learning through computer-mediated communication. It is important to understand all the details of a given mailing list community, to its uniqueness of the participants.

Organizational Memory Information Systems

Ackerman suggests that organizational information systems should provide, "new ways to access, maintain, and promote organization's intellectual assets, [that] can assist in employee turn-over, down sizing, and internationalization of personnel" (Ackerman, 1998, p203). An organization can be rendered more effective if knowledge of an organization's past is made accessible through information systems (Stein & Zwass, 2001). Organizational memory can appear in two forms: semantic (general) and episodic

(context-specific). Semantic includes organizational practices stored in handbooks, manuals, standard operating procedures, scientific knowledge (Stein & Zwass, 2001), as well as computer-mediated communication (Hahn & Subramani, 2000). Episodic memory includes situation decisions and the outcomes of those decisions (Stein & Zwass, 2001). An information system would be a significant asset to the organization if semantic and episodic memory could be stored and made accessible.

Organizational memory information systems should facilitate organizational learning (Hackbarth & Grover, 1999). The learning can occur when the knowledge is transferred from one unit or an individual, to another unit or individual. Such systems can facilitate learning by bridging the gap between space and time through synchronous and asynchronous communication (Goodman & Darr, 1998). Various task-specific information systems for organizational effectiveness have been established as shown in Table 1 (Stein & Zwass, 2001, p93). This table effectively cites past systems that support the capture of organizational memory for specific tasks. Table 1 displays that no systems have been developed to facilitate work-group practitioners; those with similar job functions but reside in various locations (Markus, 2001).

IS support of organizational memory has included identification and accessibility of subject experts and retrieval of relevant documentation (Ackerman, 1998, Zhu & Chen, 2001). Decision-making and problem solving is important to the retrieval capabilities of an organizational memory information system (Hackbarth & Grover, 1999). This importance is due to the fact that organizational memory used in information systems is formalized, captured, made explicit, and can be readily transferred (Hackbarth & Grover, 1999).

Table 1: Task-Specific Information Systems for Organizational Memory

Memory Type	Nature of Support	References
Group/team Memory	Small business team supported across time and projects	(Nunamaker, 1991) (Morrison, 1993)
Design rationale/ discussion memory	Preserves the evolution of product design or policy discussion in the making	(Conklin & Begelman, 1988) (Konda et al, 1991) (Klein, 1993) (Reddy et al, 1993)
Project Memory	Support of a large project, usually with distributed participants and long duration	(Lynch & Chen, 1992)
Meeting Memory	Provides continuity to a series of meetings.	(Sandoe et al, 1991)
Topical Memory	Accumulates answers on a targeted range of topics	(Ackerman & Malone, 1990) (Ackerman, 1992)
Document Memory	Provides access to a targeted set of richly described documents	(Huhns et al, 1989)
Environmental Memory	Assist in sense-making in interacting with the organization's environment	(Elofson & Konsynski, 1991) (Mascarebhas, 1989)

Note. Modified From "Actualizing Organization Memory with Information Systems" (p. 93), by E. W. Stein and V. Zwass, *Information Systems Research.* Copyright © 2001, Hanover, PA: Institute for Operations Research and the Management Sciences.

Computer Mediated Communication

Organizational memory embedded within computer-mediated communication has been shown to be a valuable source that captures knowledge dispersed amongst the participants of this type of communication (Hahn & Subramani, 2000). This insight

provides reasoning to store such communication for reuse, since it computer-mediated communication (CMC) captures dispersed knowledge. CMC is also viewed as an electronic forum where members can communicate freely with peers of a particular community (Zhu & Chen, 2001). The informality allows for the free flow of ideas to peers that may not be found through face-to-face communication. CMC can be in the form of a mailing group, mailing list (listserv), newsgroup, discussion list, bulletin board, etc. (Xiong & Donath, 1999). Since CMC comes in many forms, special requirements would have to be established to effectively reuse and analyze such communication. Collaborative computer-mediated communication has also included the use of group support systems (GSS) to mediate the meeting discussions of team members (Nunamaker et al, 2001). Group support systems provide a more structured environment to CMC, as Hahn and Subramani state most electronic communication is free and unstructured (2000).

CMC, as shown in this study, supports the use of peers offering suggestions to job-related problems or issues, followed by peer review of those suggestions in subsequent responses. A person who posts more answers and participates more than other users may be regarded as an expert in that area (Ahuja & Carley, 1998). Specialists and management may want to know a topic expert to ask an ad hoc question. The volume of messages sent in a CMC community has been related to the attitudes of the participants towards the community (Sproull & Kiesler, 1991). Those actively participate in a CMC process may feel a heightened sense of community with their peers, which opens a willingness to readily share knowledge.

The Communication-Garden system, developed at the University of Arizona Artificial Intelligence (AI) Lab, was shown successful in organizing the content of computer-mediated communication, in addition to visualizing the behavior patterns of the participants (Zhu & Chen, 2000). The CMC utilized by the Communication-Garden System is newsgroup communication, also known as threaded discussion, as shown in Figure 2 (Subramani & Hahn, 2000). The OrgDiscovery system, developed for this study, also includes the ability to index content and visualize communication (Grayson et al, 2002) but uses mailing list (listserv) communication, also known as Freeform, as shown in Figure 3 (Subramani & Hahn, 2000).

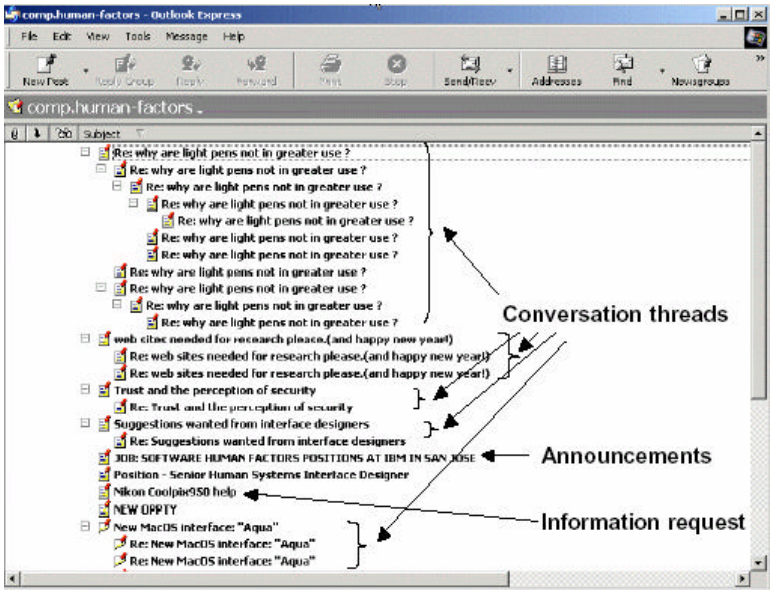


Figure 2: Threaded Discussion of a Newsgroup

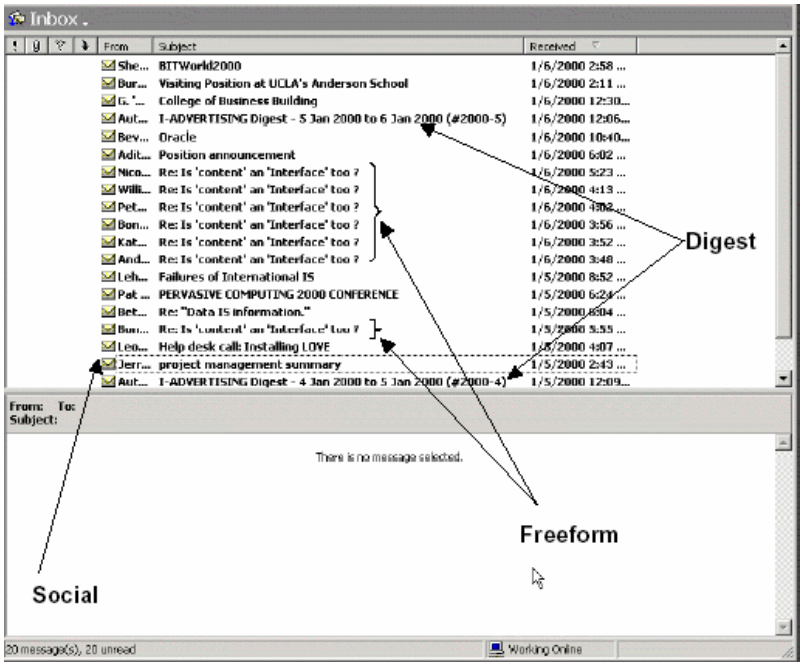


Figure 3: Freeform Discussion of a Mailing List

Note. Figure 2 and Figure 3 are modified From "Examining the Effectiveness of Group Communication Technologies: The Role of the Conversation Interface" by M.R. Subramani and J. Hahn, presented at the Academy of Mgmt Conference, http://www.jungpil.com.

Both interfaces, threaded and freeform, are given ratings of "high" for (1) visibility of social cues, (2) salience of social cues, (3) visibility of content cues, (4) and salience of content cues. The flexibility of conversion is rated "high" with freeform and "medium" with threaded. As seen with Figure 3 and Figure 4, threaded is rated as "high" for organization of conversation and freeform rates as "low" (Subramani & Hahn, 2000). Since such communication, mailing lists, newsgroups, etc., is rich in social indicators,

there has been a need for more technology to enhance the understanding of social and behavioral patterns (Zhu and Chen, 2001).

There are various scenarios where individuals within an organization transfer knowledge through computer-mediated communication, given a particular context in which that knowledge is needed. Markus (2001) states various situations and work environments where knowledge is reused. This classification includes: (a) shared work producers, various members whom are part of a team, heterogeneous or cross functional, (b) shared work practitioners, people doing the same work but in different settings, (c) expertise seeking novices, people with an occasional need for expert knowledge, and (d) secondary knowledge miners, people who seek to answer new questions through analysis of records (repository).

(Grayson et al, 2002) extrapolated from this framework to propose similar scenarios, as shown in Table 2, to portray knowledge sharing, reuse, and a search of computer-mediated communication (CMC). It includes managers who aid and assist shared work producers and shared work practitioners. These managers have a need to be "aware" of conversations that publicly take place through CMC (Grayson et al, 2002). Having the understanding of how knowledge is shared and transferred given a particular context, will allow for the development of an information system that could capture, index, and analyze the computer-mediated communication found with the mailing lists particular to this study. This analysis could provide the organizational knowledge necessary for management and mailing list members alike.

Table 2: Knowledge Sharing and Transfer via Computer-Mediated Communication

	Shared Work Producers	Managers (Shared Work Producers)	Shared Work Practitioners	Managers (Shared Work Practitioners)	Expert/ Knowledge Seekers
Description	Individuals on the same team working towards a completion of task: homogeneous or cross-functional	Managers who are responsible for the successful completion of a project/task	Individuals on different teams, doing the same work in different settings.	Provides training, documentation, techniques/tools to individuals in the same career field or job-function	Individuals who on occasion, are seeking topic/subject experts or documents of knowledge
Purpose of Desired Knowledge	Produce knowledge for their own re-use, e.g. project details, decisions made, etc.	Collective view and understanding of all issues regarding the completion of the project/task at hand	Acquire knowledge that is unknown by the individual or team. Advice could be task- specific.	Acquire knowledge of strengths and weakness of these individuals to improve training, documentation, techniques/tools	Need an answer to a sudden problem via human advice or ability to retrieve documentation
Computer-Mediated Communication (CMC) Forum	Mailing group, where shared work producers communicate project or task specific issues	Mailing group, where manager is "listening" to communication of the group for awareness of needs, status, and problems facing the team	Career-related listserv, bulletin board, etc., which provides a forum for questions and answers with peers having similar job specifications	Career-related listserv or bulletin board, where manager is "listening" to the communication for awareness purposes	Processed CMC that identifies subject/ topic experts and has organized and indexed narratives of CMC by topic

Note. From "Air Force Organizational Memory Information System: A Computer-Mediated Communication-Perspective" by M.A. Grayson, D.P. Biros, M. Ward, H. Chen. *Proceedings of the IASTED International Conference on Information and Knowledge Sharing (IKS '02),* US Virgin Islands, Nov 2002, 54-60.

Table 2 provides an overview of how task-oriented knowledge can be transferred, shared and reused via CMC given a specific context. People who need answer to an ad hoc question, may need to be referred to an expert, or have ability to retrieve relevant documentation. This table notes that managers of shared work producers (team

members) and shared work practitioners (specialists) may not be active participants in CMC. Rather these managers have a invested interest in patterns of content and behavior to further manage such employees.

Conversation Visualization

In a seminal 1987 paper, McCormick, DeFanti and Brown define visualization as a "…method for seeing the unseen, for transforming the symbolic into the geometric, for generating images and interpreting images, and more" (1987, p2). Collins refers to visualization as transformation and analysis that aids in the formation of a mental picture of symbolic data. Such a picture is simple, persistent, and complete (1993). Information visualization is about utilizing interactive graphics by leveraging constraints of our cognitive, perceptual, and motor systems (Roa & Sprague, 1998).

Many visualization techniques have been created to display graphical representations of behavioral patterns found within a CMC process. Chat Circles is a graphical interface for synchronous, real-time, communication (Donath, et al 1999). As shown in Figure 4, each participant is represented as a colored circle on the screen in which his or her words appear. The circles grow and brighten with each message, and then fade and diminish in periods of silence. These circles do not diminish if a member is still connected to a chat room. One can view who is active while watching the emergence and dissolution of conversational groups. The dynamic visualization is deemed appropriate for synchronous (real-time) communication. It is useful in visualizing behavioral patterns in real-time, but lacks the identification of whom a circle

represents. It cannot make conclusions about content or behavioral patterns when a chat room session is complete.

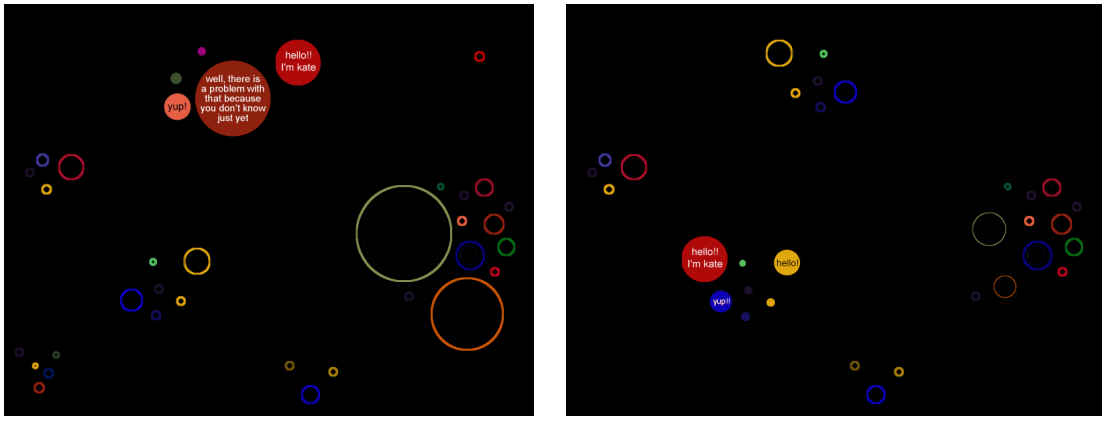


Figure 4: Two Frames from a Chat Circle Session

Note. Modified From "Visualizing Conversation" by J. Donath, K. Karahalios, & F Viégas, Journal of Computer-Mediated Communication 4(4), http://www.ascusc.org/jcmc/vol4/issue4/donath.html.

Loom is a visualization tool for threaded newsgroup discussion (Donath, et al 1999). It creates a graphical representation of participants and interactions in a threaded newsgroup. The patterns and texture of the events within the group are reflected in the patterns and texture of this digital fabric. It provides a visual interface for browsing the newsgroup archives to help users perceive social patterns. Loom also traces the connections between sequential posts in a given thread as shown in Figure 5. Lines connect the communication thread as it connects from person to person. Again we see that people are not identified explicit by name, rather represented as a document. If a

user wanted to identify experts on topics or gain more information on individuals, this graphical representation does not have the capability. It does have the ability to provide summaries of the behavior patterns of the group, not the content.

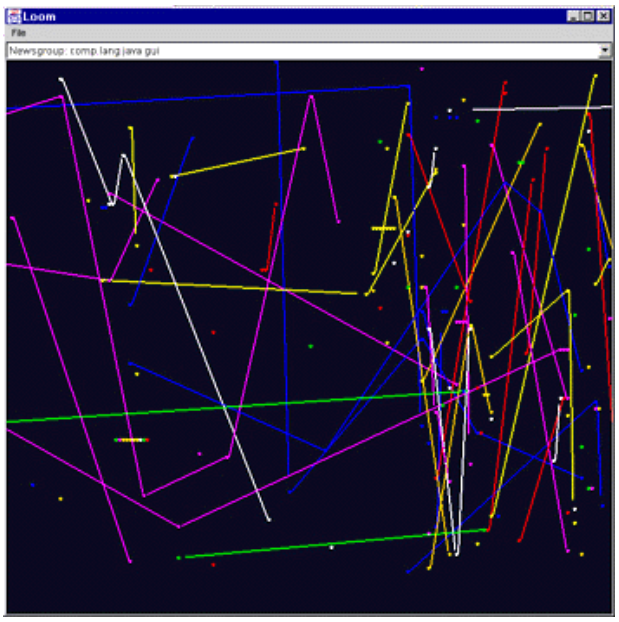


Figure 5: Loom Showing Connections between Postings in Same Thread

Note. Modified From "Visualizing Conversation" by J. Donath, K. Karahalios, & F Viégas, Journal of Computer-Mediated Communication 4(4), http://www.ascusc.org/jcmc/vol4/issue4/donath.html.

The PeopleGarden is a system that provides a graphical overview of discussion group conversation (Xiong & Donath, 1999). This system creates a "data portrait of user conversation as shown in Figure 6. Each user on the participant is represented as a flower, called a PeopleFlower. A given petal can be used to represent different attributes about each posting: time of the posting, the amount of the response, and whether a post starts a new conversation. These three attributes are seen as the most valuable in

conveying social information about the user. The number of petals increased (flower opens) as more messages are posted. Each petal fades in color showing time in posting. Saturation of adjacent petals denotes a gap in posting. A PeopleFlower can display not only a user's own posting, but also the amount of feedback from other users of the board.

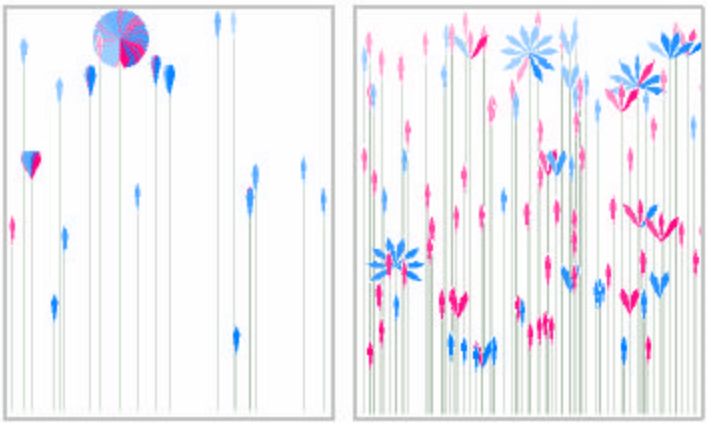


Figure 6: Data Portraits of the PeopleGarden

Note. Modified From "PeopleGarden: Creating Data Portraits for Users" by R. Xiong & J. Donath, *Proceedings of the 12th ACM Symposium on User Interface Software and Technology*, Copyright © 1999, Ashville, NC: Association of Computing Machinery.

The PeopleGarden is affective in providing a study of behavioral patterns of the group through various attributes of the flower. As shown in Figure 6, the leftmost image represents a dominating voice and the rightmost image represents a more democratic community (Xiong and Donath, 1999). This system is also capable of displaying behavioral patterns on the participants, but does not discuss anything about the content of the discussions.

The Communication-Garden system (Zhu & Chen, 2001) is inspired by the flower representation of the PeopleGarden (Xiong & Donath, 1999). The Communication-Garden system provides resources for social visualization of behavioral patterns of a CMC process as shown in Figure 7. It also includes information analysis technologies such as the Arizona Noun Phraser (AZNP) (Tolle & Chen, 2000) and a self-organizing map (SOM) (Kohonen, 1995).

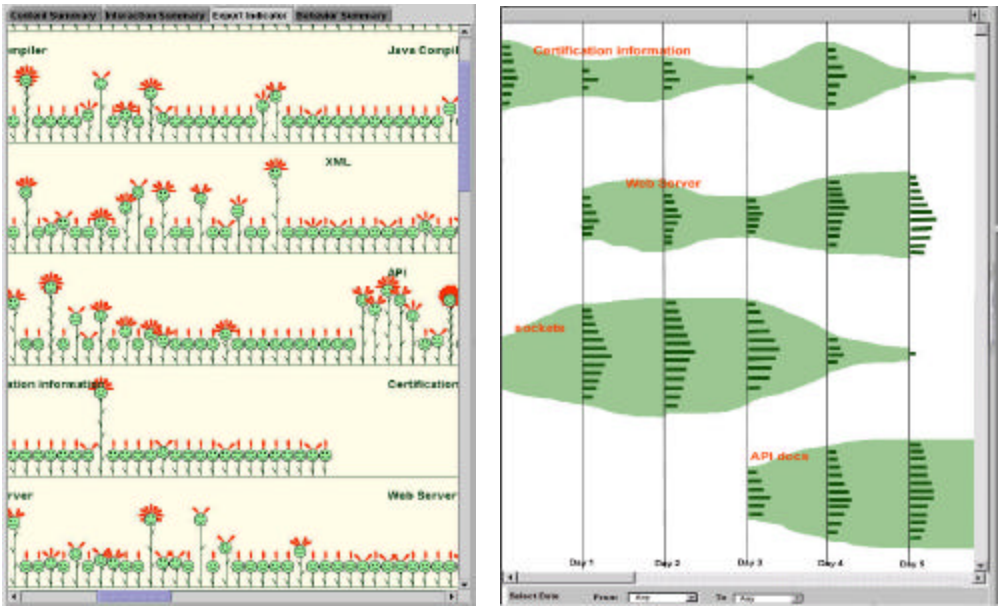


Figure 7: Communication-Garden System: Expert Indicator and Content Summary

Note. Modified From "Social Visualization for Computer-Mediated Communication: A Knowledge Management Perspective" by B.Zhu, & H. Chen, *Proceedings of the 11th Workshop On Information Technologies And Systems*, 2001, New Orleans, LA, December 16-17.

The three main components include an *Interaction Summary*, *Expert Indicator*, and the *Content Summary*. The Interaction Summary depicts the liveliness of discussion within each sub-topic by employing flower representation, as shown in the rightmost image of Figure 7. The Content Summary, depicted in the leftmost image of Figure 7, describes the temporal change in each sub-topic. The Expert Indicator uses flowers to help users locate persons active in each topic. This system is effective in identifying behavioral and content patterns through visualization, but users were shown to be overload with information (Zhu and Chen, 2001). In addition, this system represents people as flowers, but does provide information on whom the flower represents. Chat Circles, Loom, PeopleGarden system, and the Communication-Garden system, are effectively in visualizing behavioral patterns. The Communication-Garden is effective in visualizing patterns in content as well. All systems lack the ability to pin-point specific experts on topics, such as identifying names. Such systems represent people visually, but do not identify illustrated people explicitly by name.

Natural Language Processing (NLP)

With emails, it is assumed that the subject heading of email should represent the content of the narrative within the email. Many of times, this is not the case, i.e. a subject heading that states "I have a question", does not explain anything about the content of the question being asked (Grayson et al, 2002). Tolle and Chen give an example of keywords and/or thesaurus headings that accompany an academic paper (2000). "These keywords or phrases are intended to represent the dominant topics discussed in the article…frequently this is not the case" (Tolle, p355, 2000). This is true for email

documents, where the subject fails to identify the content within (Grayson et al, 2002). Thus we cannot solely index an email based on its subject heading for in depth retrieval of relevant narratives. "Manually indexing documents can be difficult and time consuming" (Tolle, p355, 2000). Even though there could be a wealth organizational knowledge stored in each email, "analyzing huge amounts of textual data requires a tremendous amount of work in reading all of the text and organizing content" (Nasukawa & Nagano, 2001).

The purpose of Natural Language Processing is to provide human thinking to a computer for the analysis of textual language. "Computational Linguistics tries to implement this process efficiently by subdividing the task into syntax and semantics" (Tolle & Chen, 2000, p356). Natural Language Processing applies syntactic and semantic rules to label parts-of-speech and identify concept phrases. Noun Phrasing is a form of natural language processing that extracts a rich representation of a document's content (Tolle & Chen, 2000). Noun Phrasing, along with information analysis techniques, has been used organize documents and content for search and retrieval of relevant documentation. Such a process that can be performed to extract meaningful phrases using this approach should include:

1) **Tokenizing** is necessary to break the stream of characters and punctuation into discrete words and sentences.
2) **Stemming** identifies the base form of each word in the text.
3) **Tagging** identifies the part-of-speech for each word.
4) **Phrase extraction** identifies unique and important concepts that often appear as multi-word phrases within the narrative. (Roa & Sprague, 1998).

Figure 8, a syntax parsing tree, demonstrates the linguistic approach to phrase extraction.

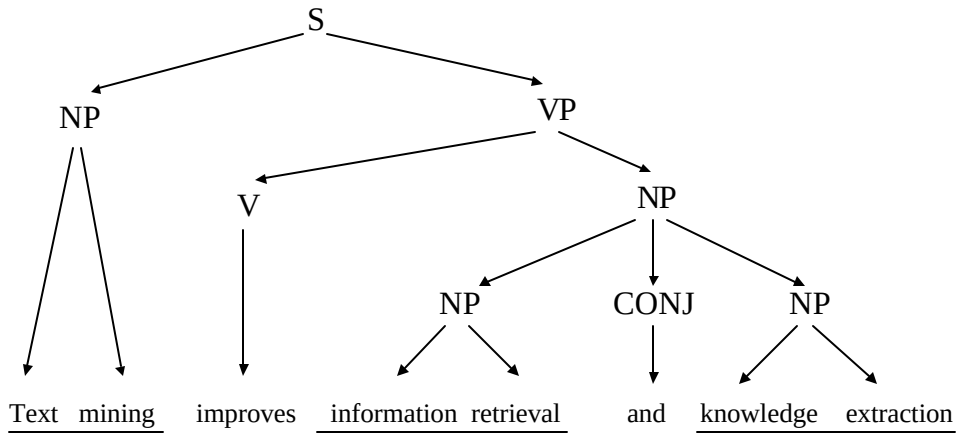


Figure 8: Syntax Parsing Tree Diagram

The Language Technology Group at the University of Edinburgh has developed LT Chunk, which is a syntactic or partial parser (Finch & Mikheev, 1997). This tool has been successful in extracting noun phrases from documents. It uses the part of speech-of-speech tagging provided by LT POS and employs mildly context-sensitive grammars to detect boundaries of syntactic groups (Finch & Mikheev, 1997). This software can currently recognize boundaries of simple noun and verb phrases. The architecture part-of-speech tagger follows three major components: a tokenizer, a morphological classifier and a morphological disambiguator (Finch & Mikheev, 1997). This tool is written in the perl programming language, which has been determined to not be easily configurable in the programming environment prescribed for this study.

Satoshi Sekine developed the Apple Pie Parser at the New York University in 1995. This parser has the ability to detect noun phrase boundaries, but does not fully

extract the phrases within its output. The parser is a bottom-up probabilistic parser that finds the parse tree with the best score by a best-first search algorithm (Sekine & Grishman, 1995). This parser tags words with Penn Tree Bank parts-of-speech tags developed at the University of Pennsylvania. The parser includes the use of a modified version of the Wall Street Journal Corpus, which includes over 8,000 lexicalized entries. The Apple Pie Parser generates a syntactic tree, similar to the PennTreeBank (PTB) bracketing system (Sekine & Grishman, 1995). This tool provides a sentence with noun phrases surrounded with brackets. Programmers using the Apple Pie Parser must provide additional work to capture the noun phrases from the documents even after the narrative has been processed

The University of Arizona's Artificial Intelligence (AI) Lab developed the AZ Noun Phraser to "extract high-quality phrases from textual data" (Tolle, 2000). This noun phrase extractor was shown high results for information retrieval purposes, recall and precision (Tolle, 2000), and its output format is easy to manipulate. A tokenizer, tagger, and noun phrase generator make up its three components. It takes text as raw input and the output conforms to the PennTreeBank word tokenization rules. Its use was highlighted in the indexing of medical documents for digital libraries, which included the implementation of the Wall Street Journal Corpus, the Brown Corpus, and finally the SPECIALIST lexicon from the National Library of Medicine for improved accuracy (Tolle, 2000). It not only extracts phrases from narratives in its output, but provides additional information like the (1) document the phrase was found, (2) number of occurrences with that document, (3) how many words make up the phrase, and (4) where

in the document that phrase is located. In examining the three noun phrase programs, the AZ Noun Phraser appears to be most suitable for this study.

Knowledge Discovery in Databases (KDD)

Since communication via computer-mediated communication (CMC) comes in large volumes, there is a need to preprocess the data to extract the organizational knowledge stated previously. Chait notes that knowledge management systems must cleanse data for successful retrieval (1998). Knowledge Discovery in Databases (KDD), as noted by Bruha, expands on the notion that "since data is collected and stored at a very large acceleration these days, there has been an urgent need for a new generation of robust software packages to extract useful information or knowledge from large volumes" (Bruha, 2000, p363). There are many attributes that can be identified with an email: sender (author), subject, creation-time, number of attachments, cc name, bcc name, as well as the narrative. Bruha also notes that preprocessing data is facilitated by "…selecting and ordering attributes (features) according to their informativity" (Bruha, 2000, p364). For example, the number of attachments does not give the required organizational knowledge that management may hope to achieve.

The sender name, included with an email in United States Air Force, contains additional information that extends past the name of the sender. Air Force Instruction 33-119, Section 5.3, standardizes the convention of a sender name displayed in an email (United States Air Force, 1999). The Sender name is display as [name, rank, unit/location]. Since the information processing needs of individuals, due to role and status within an organization, result in different email patterns (Ahuja, 1998), it is

important to isolate “rank”, since it explicitly denotes status. Air Force Instruction 33-119, Section 5.3, notes that rank and location/unit should appear in every sender name that appears in an email.

Technology Acceptance Model (TAM)

Davies describes the constructs of *Perceived Ease of Use* and *Perceived Usefulness* as determiners for acceptance of information technology (1989). The TAM model was developed to explain computer usage. The goal of TAM is “to provide an explanation of the determinants of computer acceptance that is general, capable of explaining user behavior across a broad range of end-user computing technologies and user populations, while at the same time being both parsimonious and theoretically justified” (Davies, et al, p985). TAM represents an important contribution to the IS community in understanding behaviors in accepting and using technology. Figure 9 gives an overview of the Technology Acceptance Model (TAM).

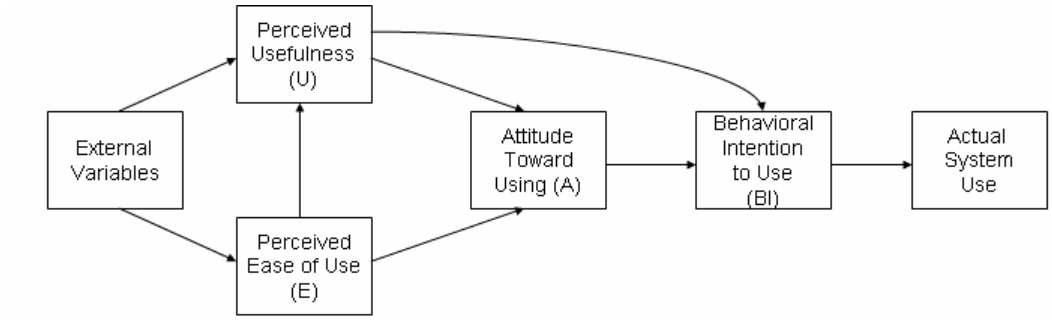


Figure 9: Technology Acceptance Model (TAM)

Based on Davies, et al, 1989

TAM notes external factors on internal beliefs to technology acceptance, as well as attitudes and intentions (Davies et al, 1989). This model’s acceptance is largely due to its generalizability to the IS community. This model views *perceived ease of use* and

perceived usefulness as the primary determiners for user acceptance. *Perceived usefulness* is a user's subjective probability that using a specific application will increase job performance. Perceived ease of use is the degree a user expects the system to be free of effort. TAM achieves these goals by identifying a small number of variables suggested by previous research that deal with the cognitive and affective determinants of computer acceptance (Davies et al, 1989). Thus, TAM is an important measure of usability in terms of increasing an individual's performance in an organizational setting.

The Technology Acceptance Model (TAM) has been implemented through the use of survey instruments to measure acceptance of information technology (Davies, 1989). This instrument provides valid measurement for predicting user acceptance of technology and has been validated through factor analysis of the constructs *perceived ease of use* and *perceived usefulness*. This instrument provided a validated instrument where industry was using invalidated surveys for development, implementation and evaluation of new products. The experiment consisted of users employing an email application for daily communication with peers. The resulting survey instrument consists of 12 subjective items on a 7 point likert-scale: 6 items that measure the construct *perceived usefulness* and 6 items that measure the construct *perceived ease of use*.

The TAM model has since been validated in studies for evaluating information technology acceptance. Adams, et al, evaluated three popular software programs, WordPerfect, Lotus 1-2-3, and Harvard Graphics (1992). The results demonstrated reliable and valid scales for the measurements of *perceived ease of use* and *usefulness.* Szajna validated this instrument through user evaluation of prepackaged database management system software (1994). Reliability for both usefulness and ease of use

scales were high with a Cronbach's coefficient of .95. Item loadings were from .77 to .92 for usefulness and .64 to .93 for ease of use. The Technology Acceptance Model (TAM) and Davies' measurement scale for *perceived usefulness* and *perceived ease of use*, has been proven as a valid instrument for measure of technology acceptance. The TAM model can be seen as an important component to the overall *usability* of any product created to improve user job performance.

Doll and Torkzadeh (1988) developed a survey instrument that measured various constructs of end-user computing satisfaction. *Perceived ease of use* in their survey consisted of 2 items that measured this construct. Zhu and Chen (2000) later used these in their evaluation of the Communication-Garden System. For this study, it was noted that Doll and Torkzadeh's 2 items for *perceived ease of use* were more applicable to this study based on the wording of the items as well as their high item loadings of .85 (1988). The 6 items for *perceived usefulness* of Davies, et al (1989) were selected for this study to the applicability and wording of the 6 items for the two software programs evaluated.

Dumas and Redish view usability as the "…means that the people who use the product can do so quickly and easily to accomplish their own tasks" (1993, p4). Usability issues can be thought of as how easy a product is to use (Jordan, 1998). For this study it is important to show why a software application is worthy of use. User-friendliness is a major component in defining the success of a product, especially for software endeavors. Usability pertains, not only to the look and feel, but also the satisfaction of the user (Shneiderman, 1998). Consistency in a products interface enables the user to successfully move from one situation to the next (Jordan, 1998). From a

software perspective, development should allow usability and users' needs to drive design decisions (Dumas & Redish, 1993).

The OrgDiscovery has been created as a more usable method to manage career-related computer-mediated communication (CMC). The next chapter, methodology, describes the prototype organizational memory information system, OrgDiscovery, developed to visualize both the content and behavior patterns found in a CMC process. This chapter examines the 5 hypothesis that answer the research question: Can an organizational memory information system (OMIS) be built to provide an understanding of behavioral and content patterns found in mailing list communication through visualization? Can this system be shown as more usable by end-users than the current software, Microsoft Outlook, used to identify such patterns?

Also this chapter will provide details of the usability experiment established for this study. This experiment was designed to see if the OrgDiscovery is more effective in task completion and rated higher in user-satisfaction than the current email program utilized by management.

III. Methodology

Introduction

This chapter provides overviews of the two software programs that have been evaluated in their abilities to identify patterns in content and behavior in mailing list communication. The two systems include: (1) Microsoft Outlook and (2) the OrgDiscovery system. Microsoft Outlook is currently used by Air Force management and specialists to: (a) read mailing list communication (b) store and reuse mailing list communication, and (c) make assumptions about the content and behavioral patterns of the participants. The OrgDiscovery system is the organizational memory information system (OMIS) designed specifically for this study. OrgDiscovery is intended to visualize both the content and behavior patterns of computer-mediated communication (CMC) as well as establishing components for organizing, retaining, and retrieving narratives for reuse. Although OrgDiscovery includes tools for retrieval of relevant documentation, the experiment is not an evaluation on information retrieval, rather identification of patterns in content and behavior.

This empirical study follows the similar design of evaluation between the Communication-Garden system and Netscape Manager (Zhu & Chen, 2001). This chapter takes a look at the four methods considered to evaluate usability between the two programs: (1) heuristic evaluation, (2) usability testing, (3) guidelines, and (4) cognitive walk-through. This chapter also explains the method of evaluation chosen to evaluate usability of the OrgDiscovery system and Microsoft Outlook. It further looks into the empirical study designed to measure the usability of both systems for a given set of tasks.

The research questions are revisited to explain the purpose behind the experiment and the results that were hoped to be achieved.

Research Approach

The method used to gather information for the investigation was based on Dumas and Redishes guide to usability testing, the method of evaluation selected to collect usability data. An overview of the steps is as follows.

1. identify user needs with managers of the career-related mailing lists,
2. observe and collect mailing list emails as an unobtrusive participant,
3. analyze previous systems built to visualize computer-mediated communication,
4. review literature on how and why to create an organization information system,
5. select specific mailing list (referred to as the “target” list),
6. analyze the characteristics of the selected mailing list
7. develop the software to visualize behavior and content patterns of the mailing list,
8. develop the research instrument,
9. select subjects who are familiar with current software used to manage email,
10. administer the survey,
11. gather survey results,
12. perform statistical analysis of the final data, and
13. interpret the results.

Interviews

Many interviews were conducted with the Air Force Communication's Agency to gain information on what knowledge they hoped to gain through implementation of a new system that would present mailing list content in a more usable fashion. Conversations took place from December 1, 2001 to August 5, 2002, as the development of the software took place. These informal discussions provided feedback based on visual "screen shots" emailed to managers of the mailing lists as the prototype system was being developed. Management decided the knowledge they wanted to gain from this system. The knowledge they hope to acquire included:

(1) Identification of subject/topic experts within the community.
(2) Organizational strengths/weaknesses - topics expressed in the form of questions and answers, and the frequency of such topics.
(3) Content/behavior study - which includes an analysis of topics in correlation to the characteristics (ranks) of the individuals.
(4) Time-sequenced events, frequency of topics given a specific time frame, and
(5) Indexed knowledge repository of discussions to provide a framework for reuse.

Based on the "look" of this system, AFCA wanted furthered usability testing to occur at the Air Force Institute of Technology (AFIT).

Observation

The observation included the non-obtrusive monitoring and collection of various AFCA mailing lists' emails. All of the mailing lists supported by AFCA are designed to support communications and information activities. The Air Force Comm and Info community is further subdivided into 4 specific skill sets, C-E maintenance (2E), information management (3A), Comm-Computer Systems (3C), and Visual Information (3V). It was important to ensure that each mailing list was used by the participants (specialists) as an open forum for questions and answers. The mailing lists, provided to specifically support C-E maintenance and various Comm-Info career fields, are listed below in Table 3. The numbers of participants, as of January 12, 2002, are also included.

Table 3: Mailing List and Number of Participants

C-E Mailing Lists		Career/Support Mailing Lists			
List	**# of People**	**List**	**# of People**	**List**	**# of People**
ANGAFRC	74	2ECHIEFS	*154*	EITRAINING	40
BISSCCTV	65	3ACHIEFS	*99*	JEFXCOMMIPT	88
COMPSECR	226	3C0X1	*144*	PLANSFLT	464
DIALTONE	155	3C0X2	*54*	TERALINK	189
DRTRADIO	469	3C1X1	*32*	TMS	186
GNDRADAR	152	3C1X2	*11*	WIN2KWG	129
MAINCNTL	212	3C2X1	*77*	WM	410
MAINMGMT	367	3CCHIEFS	*25*		
MAINSUPT	456	3CXXX	*124*		
METENAVS	211	AFPEC	*139*		
TACPASOC	74	AIRWG	*32*		
TERALINK	109	C4ISPANDCERTS	*57*		
WIREDAWG	190	CPMRS	*13*		

As of January 12, 2002

Content Analysis

In design of the OrgDiscovery system, it was important to find a method to confidently identify when a narrative was used in the form of a question (weakness) or when a narrative was used in the form of an answer (strength). A fifteen-week pilot study was conducted to show that the subject prefix of an email, i.e. 'RE:' could identify strengths/weaknesses of the IT mailing lists participants under study. A total of 670 emails were reviewed; 246 of the emails contained no subject prefix and 424 emails were threaded emails containing the subject prefix "RE:". The research questions included:

(a) Do emails without a subject prefix contain a question (weakness)?

(b) Do emails with the subject prefix 'RE:' contain an answer (strength)?

For example, does an email in this study, with a subject heading "Modem Connection" contain a question? Moreover, does a related email with a subject heading "RE: Modem Connection", contain an answer? As Walsh and Ungson explains, "…the journalist's six questions (who, what, when, where, why, and how) provide a useful way of characterizing the scope of information that may be acquired about a particular decision stimulus and organizational response". The objective here was to find a simple way for the prototype system to note if an email was in the form of a question or not, based on the inclusion or lack of, the subject prefix "RE:". The preliminary results, show a 95% confidence that emails without a subject prefix are questions, 90.04% to 95.75% of the time. Also, we can be 95% confident that emails with the subject prefix 'RE:' are answers, 93.17% to 97.31% of the time.

Target Mailing List

The mailing list reviewed extensively as the input to evaluate the OrgDiscovery and Microsoft Outlook is the DRTRADIO mailing list, Ground Radio Systems (2E1 x3). This mailing list was isolated for this study due to the high number of subscribers, 469, as noted in the early stages of creating an archive of mailing lists emails. As an unobtrusive observer, this list was particular in the variety of topics discussed, and the sporadic pattern of participation found amongst the participants each week.

OrgDiscovery System Overview

The OrgDiscovery system has been developed to visualize the content and behavioral patterns of a computer-mediated communication (CMC) process. It has been designed to provide the organizational knowledge that management hopes to achieve about the mailing list communication. The input utilized by this system are mailing lists' emails captured on a daily basis using Microsoft Outlook as a storage tool. The OrgDiscovery System is a desktop application that provides tools to index, analyze, and visualize the knowledge in publicly shared emails stored on a user's personal computer (PC). The OrgDiscovery system can take a mailing list archive and parse (cleanse) each email, and map the relevant information to a database, as shown in Figure 10. With every new project, a new database is created for a given mailing list repository. Below is an ER diagram showing the entities and attributes that can be mapped from an email object.

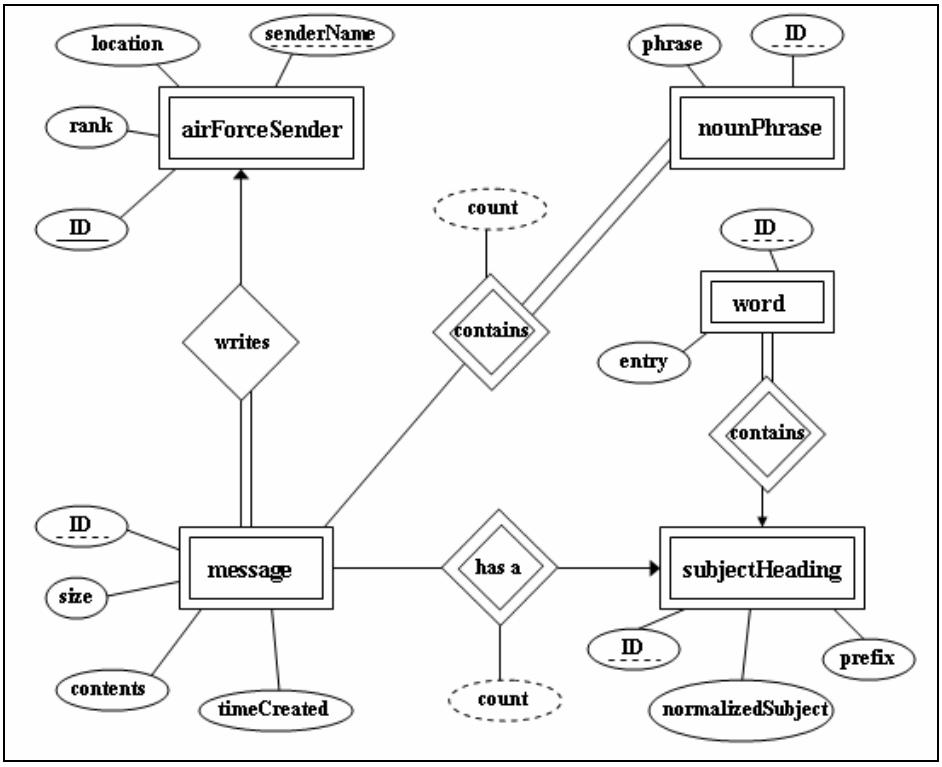


Figure 10: ER Diagram for Mapping an Email to the Database

It is important to note that the ER diagram suggests that an email must be broken down into various components. As components are extracted, items like the sender-name found in an email must be further processed. The sender name, included with an email in United States Air Force, contains additional information that extends past the name of the sender. Air Force Instruction 33-119, Section 5.3, standardizes the convention of a sender name displayed in an email (United States Air Force, 1999). The Sender name is displayed as [name, rank, unit/location]. Since the information processing needs of individuals, due to role and status within an organization, result in different email patterns (Ahuja, 1998), it is important to isolate "rank", since it explicitly denotes status.

After the mailing list emails have been processed and the content has been organized, the user can view graphical representations of the processed computer-

mediated communication through five different tabs. (1) The Database View provides a glimpse of the preprocessed data through datagrid components. (2) The Knowledge-Experts tab allows for keyword/phrase search of experts and respective narratives. (3) The Time-Sequence tab, is used to graphically display the topics discussed and respective frequencies during a given time interval. (4) The Strengths/Weaknesses tab is similar to the Knowledge-Experts tab. A user can select a unique location/unit to view (a) topics expressed in the form of a question or (b) topics (subject headings or noun phrases only) expressed in the form of an answer. Conversely, a user can also select a subject/phrase to review which units are strong or weak on that given topic. (5) The Content/Behavior tab is similar to the Time-Sequence tab, but includes the "ranks" of the participants as an additional identifier. This allows the user to view the topics discussed for a given rank (status) during a distinct time-interval. The OrgDiscovery Manual, found in Attachment A contains detailed descriptions and screen shots of the working application.

Quantitative Design Objective

The four methods considered to evaluate usability between the two programs include: (1) heuristic evaluation, (2) usability testing, (3) guidelines, and (4) cognitive walk-through (Jeffries et al, 1991). The study conducted by Jeffries et al, as shown in Table 4, noted the advantages and disadvantages to these four approaches which was taken into account in choosing the best method for this study, *usability testing*.

Table 4: Various Evaluation Methods of Usability

	Definition	Advantages	Disadvantages
Heuristic Evaluation	User interface experts review the interface and look for properties they know from experience, will cause problems	Identifies more problems Identifies more serious problems Low Cost	Requires User Interface (UI) experience Requires Several Evaluators
Usability Testing	User interface is studies under real-world or controlled conditions, with evaluators gaining data as problems arise.	Identifies serious and recurring problems Avoids low-priority problems	Requires User Interface (UI) expertise High Cost Misses Consistency Problems
Guidelines	Provides evaluators with specific recommendations about the design interface	Identifies recurring and general problems Can be used by software developer	Misses some sever problems
Cognitive Walk-through	Developers of an interface walk through tasks a typical user performs.	Helps users' goals and assumptions Can be used by software developers	Needs task definition methodology Misses general and recurring problems.

Note. Modified From User Interface Evaluation in the Real World: A Comparison of Four Techniques, by Jeffries et al, *Human factors in computing systems conference proceedings on Reaching through technology,* p.119-124.

Usability testing was selected as the method to evaluate the two applications. Evaluation of end-user satisfaction must include the subject pool performing a set of tasks using both software applications. All subjects selected are expert users in Microsoft

Outlook; usability testing will provide a platform to measure if OrgDiscovery is more usable.

Task-Set Design

To gain user perceptions of both software applications, Microsoft Outlook and the OrgDiscovery system, the subjects must use each software application by completing a given set of tasks. Dumas and Redish state that tasks for usability testing should include: (a) tasks that mimic what end-users will do with the product; (b) tasks that probe potential usability problems, and (c) tasks from the developer's experiences or concerns (1993). Table 5 provides an overview of type of tasks the subjects must complete. These five task types are the same used in the evaluation of the Communication-Garden system versus Netscape Messenger (Zhu & Chen, 2000). Appendix B contains the task-sets asked of the subjects during the experiment.

Table 5: User Tasks for Usability Testing

Task	Definition	Reference
Cluster	Covers techniques that allows user to determine whether data entries are clustered or not.	(Wehrend & Lewis, 1990; Zhou & Feiner, 1998)
Identify	Describe an object that was not known previously. Involves one attribute.	(Wehrend & Lewis, 1990; Zhou & Feiner, 1998)
Compare	User decides something based on the attribute of similar objects (emails). Involves one attribute.	(Wehrend & Lewis, 1990; Zhou & Feiner, 1998)
Correlate	If objects have multiple attributes, it should be possible to discern which objects share similar attributes.	(Wehrend & Lewis, 1990; Feiner, 1998)
Rank	Requires all subjects to browse the objects to find the extreme value. Possible for scalar and ordinal data.	(Wehrend & Lewis, 1990; Zhu & Feiner, 1998)

Instrument Development

The first phase in developing a survey instrument to measure end-user satisfaction was to generate a pool of items capitalizing on: (a) testing for usability and technology acceptance (Davies et al, 1989; Dumas & Redish, 1993; Jordan, 1998; Schneiderman, 1998); (b) empirical research that had evaluated technologies towards visualization of computer-mediated communication from the perspective of behavior (Donath, et al 1999; Xiong & Donath, 1999; Zhu & Chen, 2001), and content (Chen et al, 1998; Zhu & Chen, 2001), (c) observations of the mailing list conversations over an 8-month period, and (d) the design, implementation, and the specific workings of the OrgDiscovery system. The

constructs measured on each application, M.S. Outlook and OrgDiscovery include are shown in Table 6.

Table 6: Constructs Measured for Usability Testing

Measure	Measure Type	Tool for Measure	Reference
Effectiveness *(Extent to which goal or task is achieved)*	Objective	Task Completion *(Only one correct answer)*	(Jordan, 1998, p18)
Efficiency *(The amount of effort required to accomplish a goal)*	Objective	Time on Task *(The amount of time user needs to complete a task)*	(Jordan, 1998, p19-22)
Perceived Ease of Use	Subjective	End-User Survey	(Doll & Torkzadeh, 1988)
Perceived Usefulness	Subjective	End-User Survey	(Davis, 1989)
User Preference	Subjective	End-User Survey	(Zhu & Chen, 2001)

Effectiveness as described by Jordan is the "…extent to which a goal or task is achieved" (1998, p18). This can be measured by two methods (a) *task completion* or (b) *quality of output* (Jordan, 1998). Quality of output could be qualitatively measured based on the variable quality of output resulting in a completion of a task. Task completion, on the other hand, assumes there is only one right, distinct answer. Task completion was chosen as the method to measure *effectiveness* for this empirical study.

Efficiency is the "…amount of effort required to accomplish a goal" (Jordan, 1998, p19). This can be measured by three methods: (a) *error rate*, (b) *time on task*, (c)

or (c) *mental workload* (Jordan, 1998). Error rate considers that if a user can complete a task without making any errors along the way, the efficiency may be view as high. Mental workload is a qualitative measure of a user's perception of how difficult it was to complete a task. Time on task refers to the amount of time it takes to complete a given task. Time on task was chosen as the method to measure *efficiency* for this empirical study.

Perceived usefulness is a user's subjective probability that using a specific application will increase job performance. *Perceived ease of use* is the degree a user expects the system to be free of effort. These two constructs were measured qualitatively for each software program through an 8-item survey on a 7-likert scale. Two of the items measure *perceived ease of use* developed by Davies et al, 1989. This survey has been validated through additional empirical studies (Adams, et al, 1992; Szajna B., 1994) as has become the standard due to its generalizability to the IS community.

User preference as noted by Zhu and Chen in their empirical study of the Communication-Garden system and Netscape Messenger (2001) used a survey question asking subjects which system, did they prefer after all task-sets were completed.

Hypotheses

The research questions, as mentioned in chapter 1 includes: Can an organizational memory information system (OMIS) be built to provide a better understanding of behavioral and content patterns found in mailing list communication through visualization? Can this system be shown more usable by end-users than the current software, Microsoft Outlook, used to identify such patterns?

Based on the constructs selected for measurement in this experiment, the research questions can be narrowed to specific hypotheses for the experiment. *Effectiveness* was measured as "task completion", meaning there is one correct answer. *Efficiency*, the amount of effort to accomplish a goal, is being measured as time-on-task. *Perceived Ease of Use*, *Perceived Usefulness*, and *User Preference* are being measured in an end-user survey. The hypotheses for this experiment include:

H_1: The OrgDiscovery is more **effective** than Microsoft Outlook for the tasks performed. I.e., subjects using OrgDiscovery, answer more questions **correctly** than with Microsoft Outlook.

H_2: The OrgDiscovery is more **efficient** than Microsoft Outlook for the tasks performed. I.e. subjects answer questions more **quickly** using OrgDiscovery than with Microsoft Outlook.

H_3: Users perceive the OrgDiscovery **easier to use** than Microsoft Outlook for the tasks performed.

H_4: Users perceive the OrgDiscovery more **useful** than Microsoft Outlook for the tasks performed.

H_5: Users **prefer** the OrgDiscovery system than Microsoft Outlook for the tasks performed.

Subject Pool

The subjects used in this experiment consisted of graduate students whom are familiar with Microsoft Outlook (use daily). In terms of testing and evaluating software, it is important to select subjects who are equal in experience as the actual users (Jordan, 1998). Since the managers of the mailing lists cannot be used in the experiment, it was important to select subjects who had similar experience with Microsoft Outlook and are

familiar with sorting emails by “Sender Name”, “Time”, “Subject”, etc. Participants were given training in sorting emails by various attributes with Microsoft Outlook. Training with OrgDiscovery included a brief tutorial of each of the 4 main tools evaluated. Participants were told they could use any methods with Microsoft Outlook to complete the tasks based on their experience with the software.

Data Collection Procedures

As stated earlier in this chapter, interviews were conducted to learn management’s perspective on what the OrgDiscovery software should do. Also, an 8-month review of all of the target mailing was done to examine behavioral patterns of the participants as well as learn about the variety in content discussed. The OrgDiscovery system was built and continuously modified based off the feedback management gave from viewing visual screen-shots of the software. The OrgDiscovery system was been built to use exactly the same input, mailing list emails, as Microsoft Outlook uses. The OrgDiscovery system has been built to visualize both the content and behavior patterns of the communication.

To gather quantitative data, an email was sent to the students of the Air Force Institute of Technology (AFIT) inquiring if they would participate in the experiment. The experiment consisted of four sessions a subject would participate in. The sessions are shown in Table 7.

Table 7: Prescribed User Sessions Undertaken During Experimentation

Session	OrgDiscovery	Microsoft Outlook
Session A	Knowledge Experts Tool	Text-based interface *(group by Subject)*
Session B	Time-Sequence Tool	Text-based interface *(group by Time)*
Session C	Strengths/Weakness Tool	Text-based interface *(group by Sender or by Subject)*
Session D	Content/Behavior Tool	Text-based interface *(group by Sender, Subject or "Received" time)*

Each session consists of two task-sets. Each task-set contains 5 questions that must be completed using the specified software application. The order of the sessions a participant goes through was random. Since two tasks sets are designed for each session, these sets were be assigned to OrgDiscovery or Microsoft Outlook randomly. This ensures that no particular task set is designed exclusively for one system. Whether the user uses OrgDiscovery first or Microsoft Outlook first in a given session was random. To ensure timing techniques were standard, the test was administered to each individual one at a time, with the same administrator used for each person. This ensured timing methods were standard across all participants. The clock was stopped immediately after a participant wrote an down answer, the time was recorded, at which the clock was started to measure the next task.

Statistical Analysis

Hypothesis 1 required the average number of correct answers of subjects using the OrgDiscovery and Microsoft Outlook for each of five task types. To measure the statistical differences between these two systems, a ***pair-wise t-test*** was performed for

each task type (Cluster, Identify, Compare, Correlate, Rank) to see if the OrgDiscovery system enables users to answer more questions correctly than Microsoft Outlook. In addition, an average number of correct answers for **all tasks** were taken with OrgDiscovery and with Microsoft Outlook. A ***pair-wise t-test*** was performed to see if the OrgDiscovery system enables users, overall, to answer more questions correctly than Microsoft Outlook.

Hypothesis 2 required the average time on task of subjects using the OrgDiscovery and Microsoft Outlook. To measure the statistical difference between these two systems, a ***pair-wise t-test*** was performed for each task type (Cluster, Identify, Compare, Correlate, Rank) to see if the OrgDiscovery system enables users to answer questions more quickly than Microsoft Outlook. In addition, an average number of correct answers for **all tasks** were taken with OrgDiscovery and with Microsoft Outlook. A ***pair-wise t-test*** was performed to see if the OrgDiscovery system enables users, overall, to answer questions more quickly than Microsoft Outlook.

Hypothesis 3 and Hypothesis 4 required finding of the mean value for the constructs *perceived ease of use* and *perceived usefulness*, respectively, for both OrgDiscovery and MS Outlook. A ***pair-wise t-test*** was performed to discover the statistical difference between these to applications in terms of *perceived ease of use* and *perceived usefulness*.

Finally Hypothesis 5 was user preference between OrgDiscovery and MS. Outlook was performed using a ***Bernoulli P-Value*** where OrgDiscovery is viewed as a success.

Summary

This chapter explained the research approach and the methodology used to compare user-satisfaction of the OrgDiscovery system, built for this study, and MS Outlook. The research goal was to see if visualization helps those who conduct tasks on mailing lists archives, perform their job in a more efficient and user-friendly manner. The results and analysis of the experiment are provided in the next chapter.

IV. Results

This Chapter explores the results of the experiment outlined in Chapter 3. This chapter first outlines the analytical methods and approach used to evaluate the hypotheses and research results. Section 2 outlines the demographics of the participants of the experiment. Section 3 outlines the results and analysis of the first hypothesis, the percentage of correct answers of participants using both Microsoft Outlook and OrgDiscovery. Section 4 outlines the results and analysis of the second hypothesis, time-on-task of participants using both OrgDiscovery and Microsoft Outlook. Section 5 includes an analysis of the reliability and factor loadings of the survey used to measure *perceived ease of use* and *perceived usefulness*. Section 6 outlines the results and analysis of the third hypothesis, *perceived ease of use* of OrgDiscovery and Microsoft Outlook. Section 7 outlines the results and analysis of the fourth hypothesis, *perceived usefulness* of OrgDiscovery and Microsoft Outlook. Section 8 outlines the results and analysis of the fifth hypothesis, user preference of either OrgDiscovery or Microsoft Outlook for the tasks performed by users.

Analytical Methods (Statistics)

The Experiment utilized a set of task questions and survey questions for both OrgDiscovery and Microsoft Outlook. Each set of task questions for both OrgDiscovery and Microsoft Outlook performed by a given participant was graded for correctness, while time-on-task was recorded. This provided the ability to provide results toward there being a significant difference between the means of percent correct, and the time-

on-task. Reliability and factor analyses were conducted on all 42 respondents to confirm questionnaire reliability and factor loadings of the constructs of *perceived ease of use* and *perceived usefulness*. Once reliability and conclusive factor loadings were confirmed, mean scores for perceived ease of use and perceived usefulness were computed for OrgDiscovery and Microsoft Outlook. Computing paired differences between the two systems in terms of these two constructs, provided a platform to evaluate usability between the two software programs.

Demographics

The participants of the experiment were all company-grade officers (0-1, 0-2, 0-3), who use Microsoft Outlook on a daily basis to manage emails. Limited training was provided to the participants of the experiment on Microsoft Outlook, since it was considered that participants use this tool on a daily basis and thus did not want to influence there current trends and usage of this software program. It was noted that each participant took 50-55 minutes to partake in the experiment.

Table 8: Demographic Information of Experiment Participants

Demographic		Total # of	% of Total
Gender			
Female		4	9.5%
Male		38	90.5%
	TOTAL:	**42**	**100%**
Rank			
2dLt.		7	16.7%
1stLt.		12	28.6%
Capt.		23	54.8%
	TOTAL:	**42**	**100%**
MS Outlook (Years of Use)			
1-2		5	11.9%
3-4		8	19.0%
5-6		11	26.2%
7-8		11	26.2%
9+		7	16.7%
	TOTAL:	**42**	**100%**
Listserv Member			
Yes		0	0.0%
No		42	100%
	TOTAL	**42**	**100%**

Hypothesis 1

The OrgDiscovery is more **effective** than Microsoft Outlook for the tasks performed. I.e., subjects using OrgDiscovery, answer more questions **correctly** than using Microsoft Outlook.

H_1: $\mu_{Difference} > 0$ [i.e. $(\mu_{OrgDiscovery} - \mu_{Microsoft\ Outlook}) > 0$]
H_a: $\mu_{Difference} = 0$ [i.e. $(\mu_{OrgDiscovery} - \mu_{Microsoft\ Outlook}) = 0$]

Table 9: Means (% Correct Answers) for All Task Types

	Mean (% Correct Answers)	N	Std. Deviation
M.S. Outlook	0.8321	42	15.62195
OrgDiscovery	0.933	42	5.0757

Table 10: Paired Differences of Means (% Correct Answers)

	Paired Differences						
			95% CI of Difference				
	Mean (% Correct)	**Std. Deviation**	**Lower**	**Upper**	**t**	**df**	**Sig. (1 tailed)**
Paired Difference (*OrgDiscovery - M.S. Outlook*)	.1012	.10030	.0699	.1324	.538	41	.000

OrgDiscovery assisted users in answering task questions .1012 (10%) more on the average as shown by the paired differences. A P-Value of 0.00 infers there is a significant difference between the percent of correct answered task questions by the 42 participants using OrgDiscovery vs. M.S. Outlook. Thus you can reject the alternate and accept H_1

Table 11: Paired Differences of Means (% Correct Answers) per Task Type

	Paired Differences						
			95% CI of Difference				
	Mean (% Correct)	Std. Deviation	Lower	Upper	t	df	Sig. (1 tailed)
Cluster Tasks (*OrgDiscovery - M.S. Outlook*)	.1429	.21487	.0759	.2098	4.309	41	.000
Identify Tasks	.2440	.26184	.04040	.1625	.3256	41	.000
Compare Tasks	.0357	.17083	-.0175	0889	.355	41	.183
Correlate Tasks	.0179	5996	-.0320	.0677	.723	41	.474
Rank Tasks	.0595	.14409	.0146	.1044	2.677	41	.011

Overall, OrgDiscovery assisted the participants in answering more task questions correctly. Cluster, Identify, and Rank tasks had significant differences in means (percent

correct) and low P-Values. Compare and Correlate Tasks show low differences in means and higher P-values, making it inconclusive whether M.S. Outlook or OrgDiscovery results in more correct answers for Compare and Correlate tasks.

Hypothesis 2

The OrgDiscovery is more **efficient** than Microsoft Outlook for the tasks performed. I.e. subjects answer questions more **quickly** using OrgDiscovery than Microsoft Outlook.

H_2: $\mu_{Difference} > 0$ [i.e. $(\mu_{Microsoft\ Outlook} - \mu_{OrgDiscovery}) > 0$]
H_a: $\mu_{Difference} = 0$ [i.e. $\mu_{Microsoft\ Outlook} - \mu_{OrgDiscovery}) = 0$]

Table 12: Means (time-on-task) for All Task Types

	Mean (Seconds)	N	Std. Deviation
M.S. Outlook	55.8321	42	15.62195
OrgDiscovery	28.864	42	5.0757

Table 13: Paired Differences of Means (time-on-task) for All Task Types

	Paired Differences						
			95% CI of Difference				
	Mean (Seconds)	Std. Deviation	Lower	Upper	t	df	Sig. (1 tailed)
Paired Difference (*M.S. Outlook – OrgDiscovery*)	26.9679	15.13915	22.250	31.6855	11.54	41	.000

For all task questions, M.S. Outlook had a mean time of 55.83 seconds and OrgDiscovery had a mean time of 28.86 seconds, with OrgDiscovery being 26.97

seconds faster on the average. A P-Value of 0.00 infers there is a significant difference between the completion times of all task questions asked of the 42 participants using OrgDiscovery vs. M.S. Outlook. Thus the alternate hypothesis can be rejected and can accept H_2.

Table 14: Paired Differences of Means (time-on-task) for per Task Types

	Paired Differences						
			95% CI of Difference				
	Mean (Seconds)	Std. Deviation	Lower	Upper	t	df	Sig. (1 tailed)
Cluster Tasks (*M.S. Outlook – OrgDiscovery*)	29.9643	24.12301	22.447	37.4815	8.050	41	.000
Identify Tasks	34.2083	27.52039	25.632	42.7843	8.056	41	.000
Compare Tasks	21.0476	18.68507	15.224	26.8703	7.300	41	.000
Correlate Tasks	25.6548	2.92168	19.754	31.5552	8.781	41	.000
Rank Tasks	23.8571	13.62679	19.610	28.1036	11.346	41	.000

For each task type, Microsoft Outlook took longer (means of the differences) and a P-Value of 0.000 for each task type shows high significance that OrgDiscovery provided a more efficient (quicker) platform for specific task questions completed.

Reliability

This section reviews the analysis of reliability of the survey used to measure perceived ease of use and perceived usefulness in terms of both OrgDiscovery and Microsoft Outlook. The Table below shows that the 2 items for Ease of Use are highly reliable, .87 and .86, for both OrgDiscovery and Microsoft Outlook respectively. The

Table also shows that the 6 items for Usefulness are highly reliable, .95 and .97, for both OrgDiscovery and Microsoft Outlook respectively.

Table 15: Reliability Analysis of Survey Instrument

	Cronbach's Alpha	N	# Items
OrgDiscovery			
Ease of Use	.87	42	2
Usefulness	.95	42	6
Microsoft Outlook			
Ease of Use	.86	42	2
Usefulness	.97	42	6

Factor Loadings

This section reviews the analysis the factor loadings of the items used for ease of use (2 items) and perceived usefulness (6 items) in terms of both OrgDiscovery and Microsoft Outlook. The items all have high loadings for the two constructs in terms of both software programs OrgDiscovery and Microsoft Outlook.

Table 16: Factor Loadings of Ease of Use and Usefulness

Scale Items	OrgDiscovery		Microsoft Outlook	
	Ease	Use	Ease	Use
EASE OF USE				
User Friendly	.894	.264	.789	.498
Easy to Use	.909	.240	.780	.515
USEFULNESS				
Work More Quickly	-.183	.693	.043	.827
Job Performance	-.062	.921	-.232	.946
Increase Productivity	-.193	.904	-.200	.948
Effectiveness	-.183	.952	-.138	.949
Makes Job Easier	-.009	.945	-.183	.957
Useful	.094	.935	-.125	.915

Hypothesis 3

Users perceive the OrgDiscovery **easier to use** than Microsoft Outlook for the tasks performed.

H_3: $\mu_{Difference} > 0$ [i.e. $(\mu_{OrgDiscovery} - \mu_{Microsoft\ Outlook}) > 0$]
H_a: $\mu_{Difference} = 0$ [i.e. $(\mu_{OrgDiscovery} - \mu_{Microsoft\ Outlook}) = 0$]

Table 17: Means (Ease of Use) for Systems Evaluated

	Mean (Ease of Use)	N	Std. Deviation
M.S. Outlook	3.369048	42	1.3022472
OrgDiscovery	6.130952	42	.563767

Table 18: Paired Differences of Means (Ease of Use) for Systems Evaluated

	Paired Differences						
			95% CI of Difference				
	Mean (Ease of Use)	Std. Deviation	Lower	Upper	t	df	Sig. (1 tailed)
Paired Difference (*M.S. Outlook – OrgDiscovery*)	2.761	1.535	.236	2.283	.66	41	.000

On a 7-point likert scale, M.S. Outlook had a mean score of 3.37 for *perceived ease of use*, while OrgDiscovery had a higher mean score of 6.13, with OrgDiscovery being 2.76 points higher on the average. A P-Value of 0.00 infers there is a significant difference between the *perceived ease of use* of the 42 participants using OrgDiscovery vs. M.S. Outlook. Thus the alternate hypothesis can be rejected and can accept H_3.

Hypothesis 4

Users perceive the OrgDiscovery more **useful** than Microsoft Outlook for the tasks performed.

H_4: $\mu_{Difference} > 0$ [i.e. $(\mu_{OrgDiscovery} - \mu_{Microsoft\ Outlook}) > 0$]
H_a: $\mu_{Difference} = 0$ [i.e. $(\mu_{OrgDiscovery} - \mu_{Microsoft\ Outlook}) = 0$]

Table 19: Means (Usefulness) for Systems Evaluated

	Mean (Usefulness)	N	Std. Deviation
M.S. Outlook	3.138889	42	.8339042
OrgDiscovery	6.257937	42	1.5586617

Table 20: Paired Differences of Means (Usefulness) for Systems Evaluated

	Paired Differences						
			95% CI of Difference				
	Mean (Usefulness)	Std. Deviation	Lower	Upper	t	df	Sig. (1 tailed)
Paired Difference (*M.S. Outlook – OrgDiscovery*)	3.119048	2.0496168	.31623	3.75775	.682	41	0.00

On a 7-point likert scale, M.S. Outlook had a mean score of 3.139 for *perceived usefulness*, while OrgDiscovery had a higher mean score of 6.258, with OrgDiscovery being 3.119048 points higher on the average. A P-Value of 0.00 infers there is a significant difference between the *perceived usefulness* of the 42 participants using OrgDiscovery vs. M.S. Outlook. Thus the alternate hypothesis can be rejected and can accept H_4.

Hypothesis 5

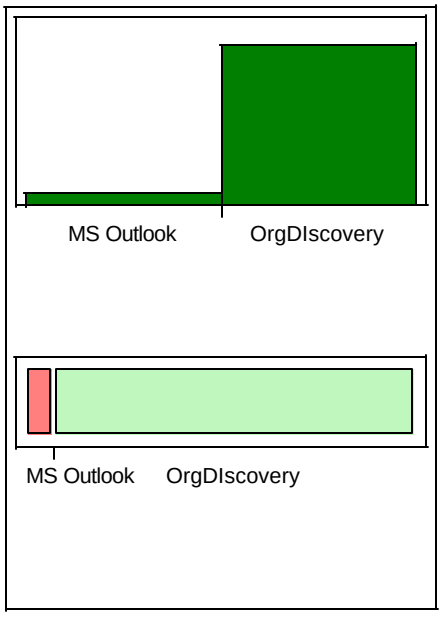
Users **prefer** the OrgDiscovery system than Microsoft Outlook for the tasks performed.

H_4: $\mu_{OrgDiscovery} > \mu_{Microsoft\ Outlook}$
H_a: $\mu_{OrgDiscovery} = \mu_{Microsoft\ Outlook}$

This test utilizes the ***Bernoulli P-Value***, where OrgDiscovery is viewed as the success. The alternate hypothesis would assume a 50/50 outcome of users selecting either OrgDiscovery or M.S. Outlook. The probability of OrgDiscovery being at 50% or less selection rate, results in a P-Value of 0.00 as shown in Table 21, showing a

significant difference in the preferred software for the tasks performed. Thus the alternate hypothesis can be rejected and can accept H_5.

Table 21: User Preference between OrgDiscovery and Microsoft Outlook



Level	Count	Prob
MS Outlook	3	0.07143
OrgDiscovery	39	0.92857
Total	42	1.00000

Level	Estim Prob	Hypoth Prob
MS Outlook	0.07143	0.50000
OrgDiscovery	0.92857	0.50000

Binomial Test	Level Tested	Hypoth Prob	p-Value
Prob <= p	OrgDiscovery	0.50000	<.0001

Conclusion

The results show that the OrgDiscovery system, for the tasks performed by the experiment participants, is more effective (higher % of correct answers), more efficient (less time-on-task), rated higher in terms of perceived ease of use and perceived usefulness, and more preferred over Microsoft Outlook for the specific tasks completed.

V. Discussion

Research Questions

Upon completing the research, the answer to the research question, "can an organizational memory information system (OMIS) be built to provide a better understanding of behavioral and content patterns found in mailing list communication through visualization," is a yes. Microsoft Outlook is a well-known tool for managing email transactions, but no studies have looked at its ability in users identifying patterns in content and behavior of mailing list (listserv) participants.

The answer to the second research question, "Can this system be shown more usable by end-users than the current software, Microsoft Outlook, used to identify such patterns," is a yes. The random design of the experiment ensured no bias toward any particular software. The paired difference analysis displayed that OrgDiscovery was more effective, more efficient, rated higher in terms of *perceived ease of use* and *perceived usefulness*, and more *preferred* over Microsoft Outlook for the specific tasks completed. Thus, the research questions, for this study, showed OrgDiscovery more usable by end-users for the specific tasks performed.

Research Question Discussion

Retention and retrieval of organizational memory has been the concentration of many conceptualized models of an organizational memory information system (OMIS). This thesis presented an extended view for system development of an OMIS from a knowledge management perspective. This system was designed to visualize the content

and behavior patterns of computer-mediated communication of Air Force IT Specialists. This study showed that visualization of mailing list communication provides a more usable method to make conclusions about the participants of mailing lists versus the text-based Microsoft Outlook.

Computer-mediated communication (CMC) has been viewed as rich in organization memory (Hahn & Subramani, 2000), and as a method for transferring organizational memory in large volumes (Grayson et al, 2002). Managers of employees reviewing such communication need a system to provide clues to further support organizational effectiveness. OrgDiscovery, created and evaluated in this study for such support, hopes to add to previous research towards actualized organizational memory information systems.

Limitations

A limitation of the study was the lack of including a real-world setting for experimentation. This is due to the distance in location between the researcher and the sponsoring agency, the Air Force Communications (AFCA). Not all the communication amongst specialist can be stored and analyzed. Mailing list members have the ability to respond directly to the knowledge-seeker, or reply to the whole list for everyone to view the response (Hahn & Subramani, 2000). Thus, only the capture of emails that are sent to all participants of a mailing list (listserv) could take be stored for use in this study. Figure 11 demonstrates the flow of mailing list communication as noted by the administrators of the mailing lists under study.

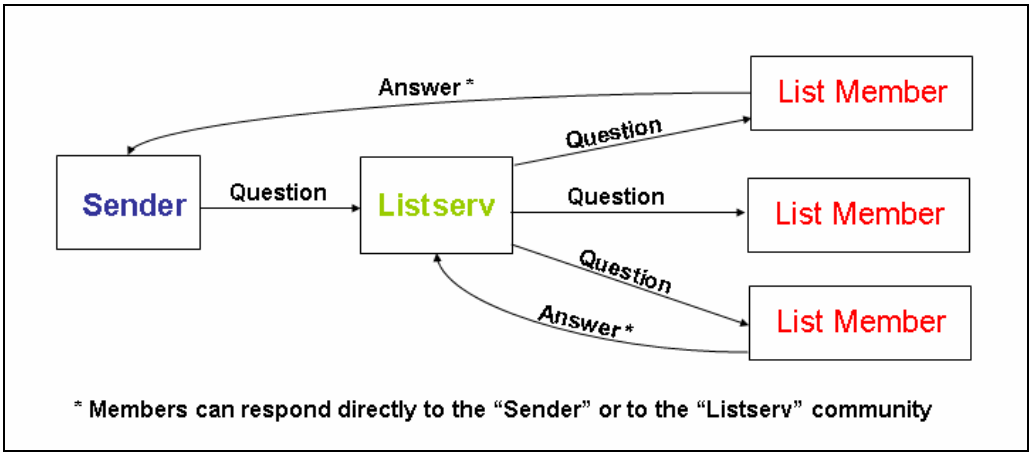


Figure 11: Flow of the Mailing List (listserv) Communication

Future Research

The OrgDiscovery system was examined in this study for its capability in assisting users to identify patterns in content and behavior. OrgDiscovery also includes the capability for information retrieval of relevant documentation as shown in Attachment A. Such future research may include information retrieval tests and evaluations on retrieving knowledge for re-use purposes.

AFCA has assumed that the best method to share knowledge amongst specialists is the use of mailing lists. This is due to messages appearing in each subscriber's inbox for notability and higher response rates to questions. Mailing lists do not provide the ability for a centrally located repository as noted by this study. Managers of these lists as well as subscribers are restricted to storing messages on their personal computer (PC). Are there other methods of computer-mediated communication (CMC) that facilitate knowledge transfer and allow for the capabilities for a centralized knowledge base of narratives for re-use purposes? This may include the usage of Usenet newsgroups. Possible research may include the application of newsgroup communication for one

particular career field using the AFCA mailing lists (listserv) as a target group. Or should the communication be synchronous, real-time communication, like with chat rooms, etc?

Further research may include social aspects of computer-mediated communication. There are many implications to monitoring communication. Are participants likely to decrease communication if monitoring methods are applied and to what extent? What group characteristics are displayed in such open forums like chat rooms, listserv's news groups etc.? These questions are important for managers wanting to supply the necessary documentation, tools, techniques, and processes to the war fighters in the field. Such social implications may include the loss of expertise when an individual leaves an organization. Is capturing the computer-mediated communication of a career-related forum an answer to such a problem? Ackerman suggests that, "new ways to access, maintain, and promote organization's intellectual assets, can assist in employee turn-over, down sizing, and internationalization of personnel" (1998, 203). What other methods can be applied to store such knowledge and expertise?

Conclusion

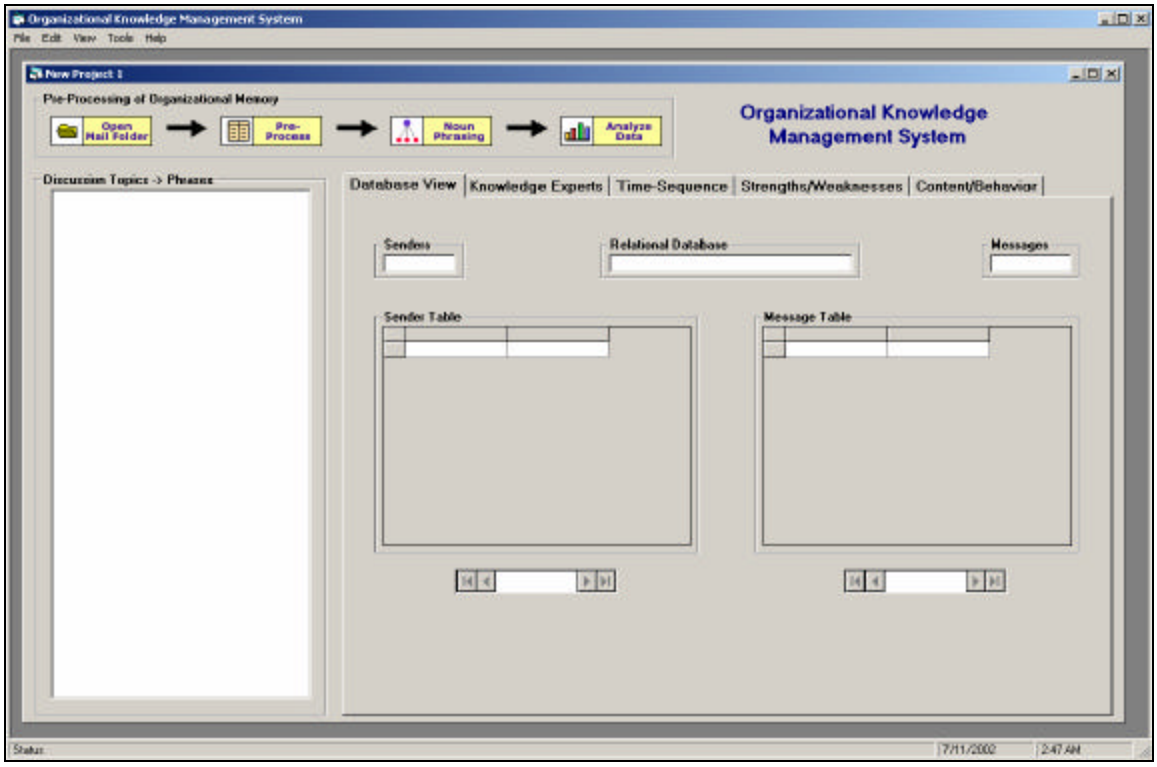
The future goal of the OrgDiscovery system is to continuously expand the functionality after continued examination of various text mining and data mining applications and methods. Capturing and analyzing the shared knowledge of the IT experts will add continuity to the high turnover rate of information technology experts in the United States Air Force. Through further implementation of the OrgDiscovery system, this on-going study hopes to provide awareness of tools and methods for capturing the knowledge of US Air Force, IT experts located around the world.

In one sense, this research appears to confirm what both academics and practitioners alike have stated about the importance of considering organizational memory when implementing knowledge management projects that assist management and worker alike. This research hopes to extend on prior studies and endeavors that have provided systems that further support organizational effectives and nurture relationships between manager and worker.

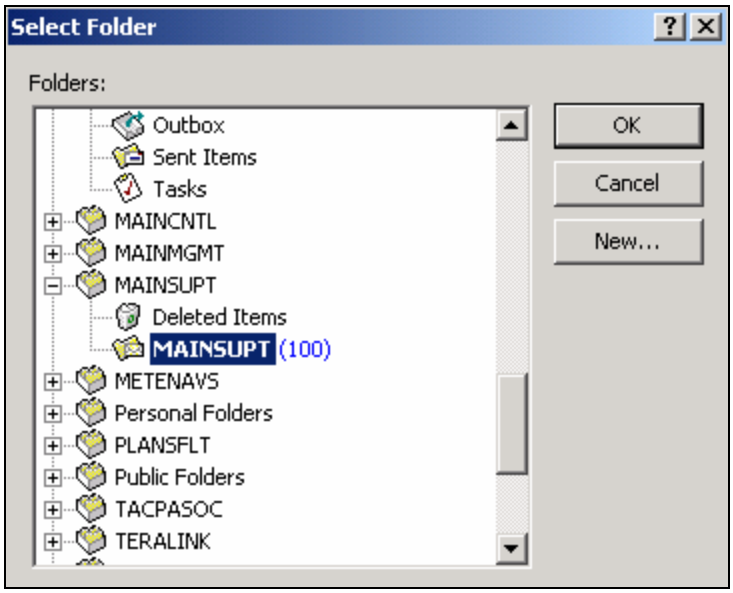
Appendix A: OrgDiscovery User Manual

The OrgDiscovery System is designed to visualize the content and behavior patterns of an archive of mailing list (listserv) communication. This manual provides an overview of OrgDiscovery's capabilities and tips for usage.

Step 1: New Project Begins



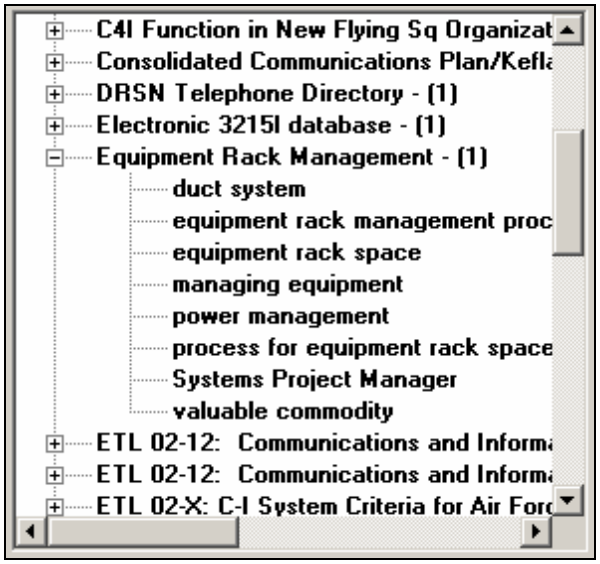
Step 2: User clicks "open mail folder" and selects a mail folder for import



Step 3: "Pre-process" button is pressed to parse email into distinct components

Step 4: "Noun Phrasing" button is pressed - extract meaningful phrases from emails

Two-Level Tree: "Subject Headings" to "Noun Phrases"

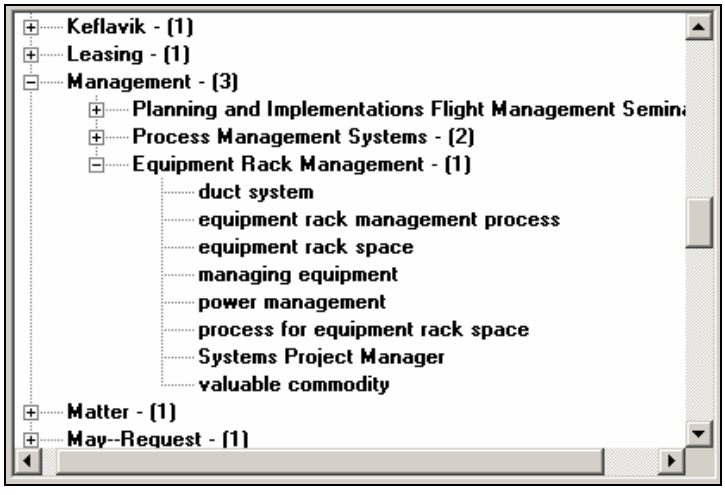


The two-level tree is a method to organize the archive of emails.

Parent Node
"Modem Connection - (3)"
Tells user there must be three emails in the repository with these respective subject headings,
"Modem Connection",
"Re: Modem Connection",
"Re: Modem Connection"

Child Node
The children nodes are the noun phrases found within the narratives of the group documents.

Three-Level Tree: "Topic" to "Subject Headings" to "Noun Phrases"



The three-level tree is a method to organize the archive of emails.

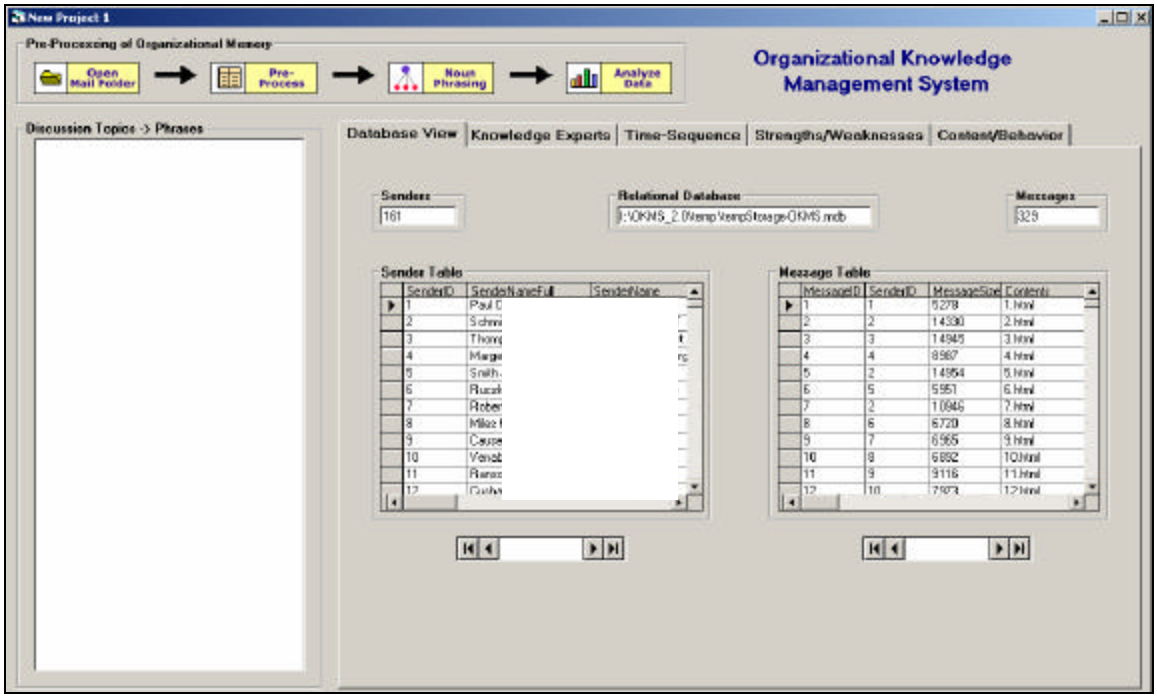
Grandparent Node
Permuted indices (topics) where their children contain the index (topic) within its text.

Parent Node
Same as 2-D Tree

Child Node
Same as 2-D Tree

Database View- (Tab 1)

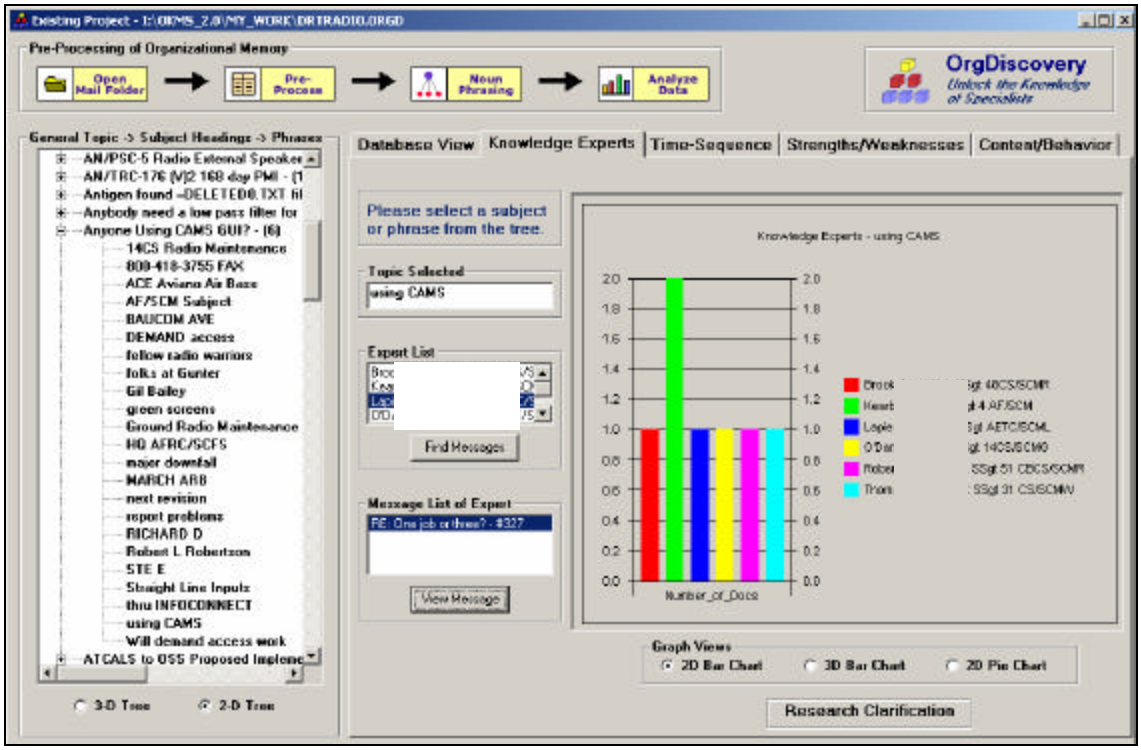
This Tab provides an in depth look at the pre-processed email messages before any Text Mining Techniques are employed. After pre-processing, the user is informed of the exact number of Senders (mailing list participants) in respect to the total number of email messages.



Preprocessed Emails Displayed in Datagrid Components

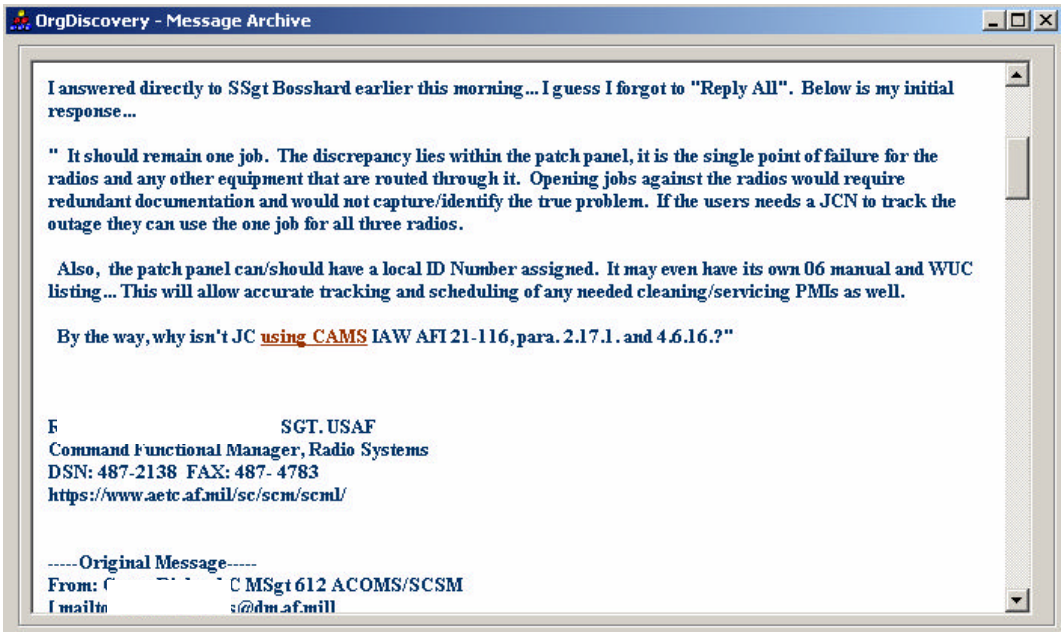
Knowledge Experts - (Tab 2)

The Knowledge Experts Tab provides the ability to search for knowledgeable specialists, given a subject heading or an internal phrase as shown in Figure 13. It will supply management (AFCA) and career-related mailing list participants, the opportunity to gain required knowledge about other specialists, e.g., identify a specialist to ask a question to an ad hoc problem or task.



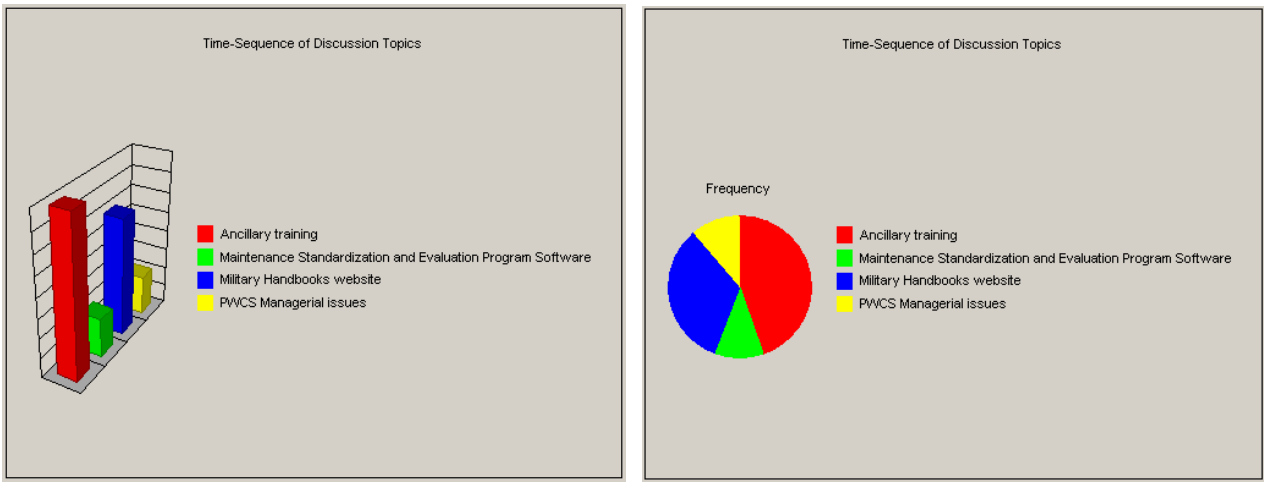
Knowledge Experts Tab

Once a user has selected an expert from the expert list provided on that topic, a list of messages appears. These messages, written by the expert, contain the selected. After "View Message" is clicked, these messages appear as shown above, and the phrase is highlighted within the message.



Message with Highlighted Phrase

3-D Bar Chart and 2D Pie Chart Representations are addition visualizations options for many of the tabs in OrgDiscovery.



3-D Bar Chart and 2-D Pie Chart Representations

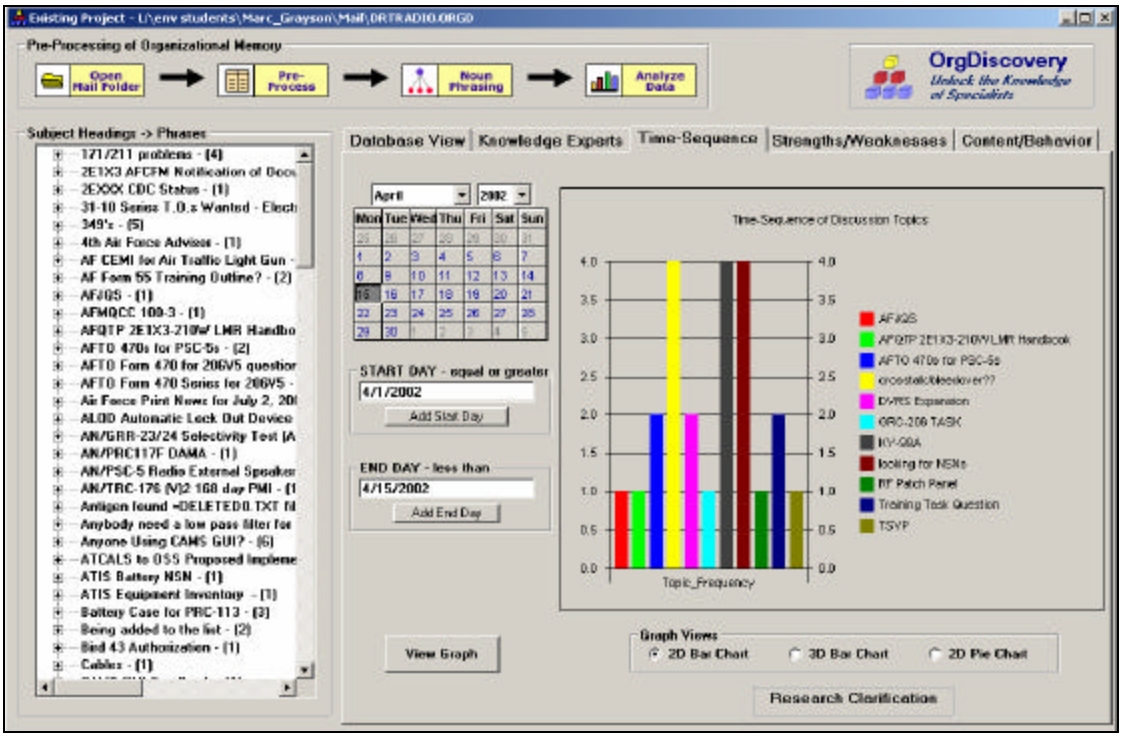
Time-Sequence - (Tab 3)

The Time-Sequence tab examines the relationship between the creation time of emails (over a distinct time-interval) and the topics/phrases found in the discussion. This view provides the capability to understand the behavior patterns of the mailing list participants based on a distinct interval. The OrgDiscovery allows user to pinpoint an exact time-frame for review, rather the entire time-span of all the communication. The steps to utilize this tab include:

(1) Select a month and a year in the calendar view

(2) Select a "start day" and "end day"

(3) Press View Graph



Time-Sequence Tab

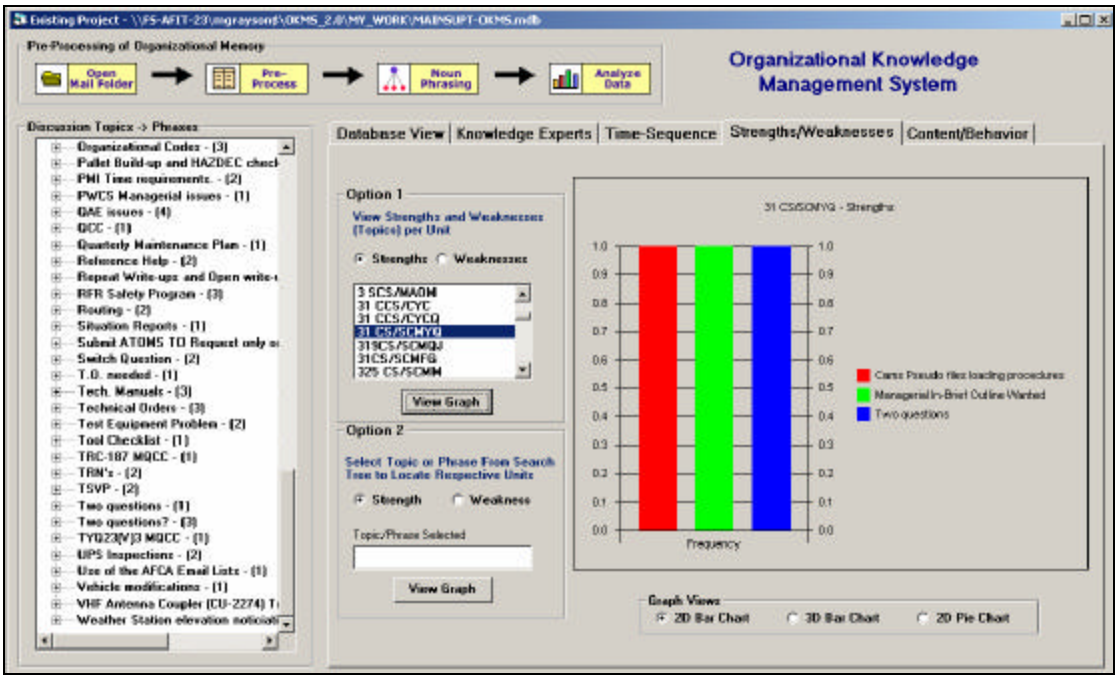
Strengths/Weaknesses - (Tab 4)

The Strengths/Weaknesses tab is similar to the Knowledge-Experts tab. Here, a user can select a unique location/unit to view (a) topics expressed in the form of a question or (b) topics expressed in the form of an answer. Conversely, a user can also select or type a topic to review which units are strong or weak on that given topic. With option 1, as shown below a user can:

(1) Select a Unit from the given list (e.g. “31 CS/SCMYQ”).

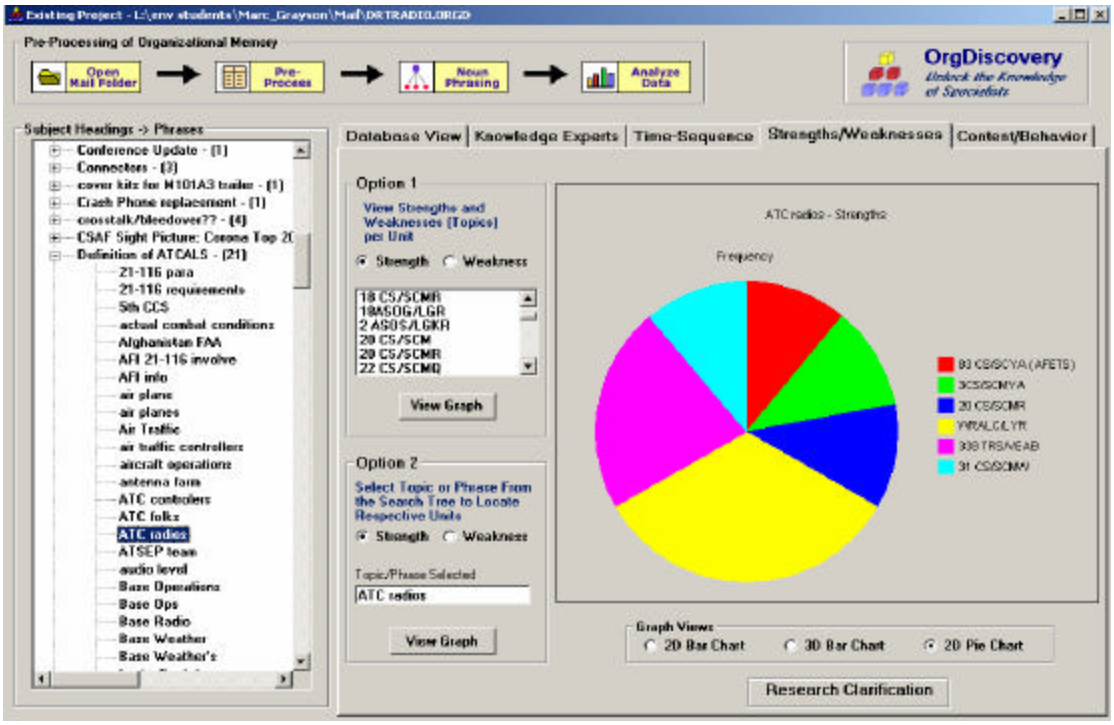
(2) Select strengths or weaknesses option.

(3) Press “View Graph” button to view the results.



Strengths/Weakness Tab (Option 1)

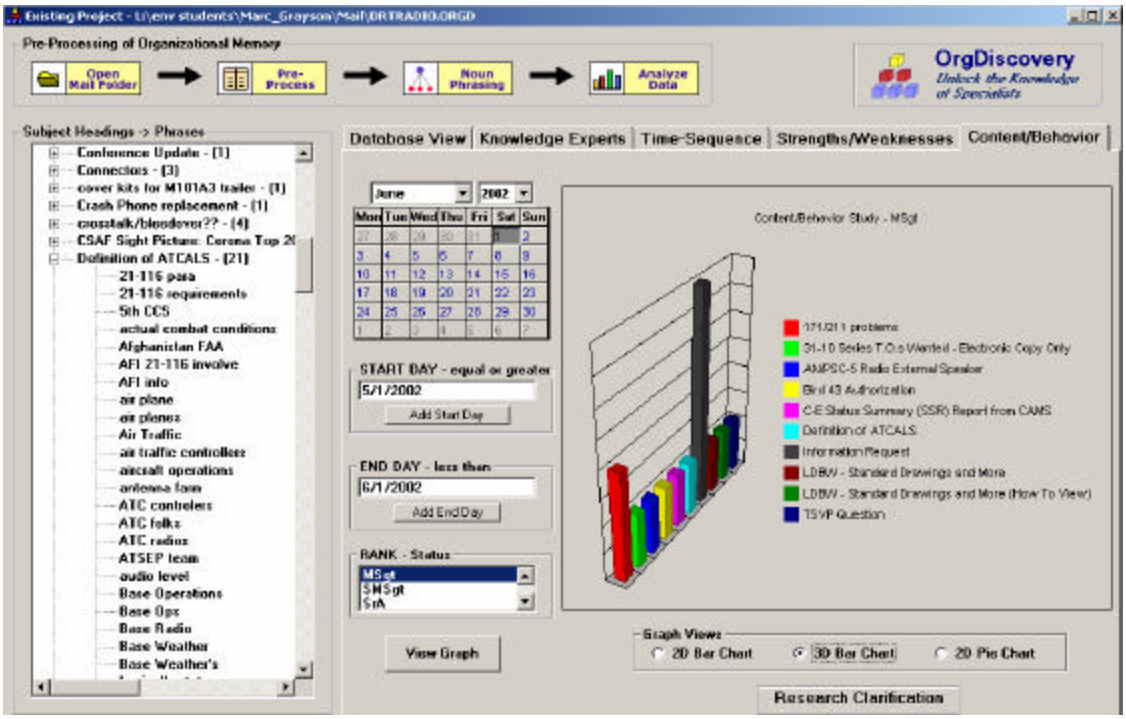
Option 2 of the Strengths/Weaknesses Tab provides a method of finding units who have persons whom are strong or weak on certain topics as shown in Figure 17. The steps to gain such knowledge by the user include selecting a topic from search tree, e.g. “ATC Radios”. A list of units is returned graphically displayed to the user.



Strengths/Weakness Tab (Option 2)

Content/Behavior - (Tab 5)

The Content/Behavior tab is similar to the Time-Sequence tab, but includes the "ranks" of the participants as an additional identifier. This allows the user to view the topics discussed for a given rank (status) during a distinct time-interval, as shown in Figure 8. Given a distinct time-interval, we can note the discussion of topics given a particular rank. The steps taken to achieve such knowledge include: (2) a user selects a month, e.g. “May”, a year is selected, e.g. “2002”, (2) a rank is selected, e.g. “MSgt”, (4) the user presses the “View Graph” button to view the organizational behaviors of a given rank during a distinct time interval.



OrgDiscovery Content/Behavior Tab

Appendix B: Usability Task Set

SESSION A	**Knowledge Experts Tool vs. Text-based interface *(group by Subject)***		
Task Set 1:	**Cluster**	Identify all persons (if any) who participated in discussions on subject "**Definition of ATCALS**" that are from the same unit?	**Answer:** NONE
	Identify	Identify the total number of **distinct** persons who have participated in discussion on the subject "**Anyone Using CAMS GUI?**"	**Answer:** 5
	Compare	Who has participated more on the subject "**AN/PSC-5 Radio External Speaker" Weiss Hans U.** or **Staley John R.**?	**Answer:** Weiss Hans U.
	Correlate	Did all the participants who responded to subject "**Gray Transit cases**" respond only once?	**Answer:** NO
	Rank	Which person has **responded** most to the subject "**Maintenance control database**"?	**Answer:** Johnson Jeffrey A.
Task Set 2:	**Cluster**	Identify all persons (if any) who participated in discussions on subject "**LMQCC for ETVS**" that are from the same unit?	**Answer:** NONE
	Identify	Identify the total number of **distinct** persons who have participated in discussion on the subject "**PRC-113 Question**"	**Answer:** 4
	Compare	Who has participated more on the subject "**Anyone Using CAMS GUI?**", **O'Daniel Richard** or **Kearby Kevin J**.	**Answer:** Kearby Kevin J.
	Correlate	Did all the participants who responded to subject "**GRM-115**" respond only once?	**Answer:** YES
	Rank	Which person has responded most to the subject "**Definition of ATCALS**"	**Answer:** Raney Douglas R.

SESSION B	Time-Sequence Tool vs. Text-based interface *(group by "Received" time)*		
Task Set 1:	**Cluster**	What subjects were discussed during the month of **January, 2002** and had the same total # of email messages?	**Answer:** NONE
	Identify	Identify the total number of distinct topics (subjects) during the month of February, 2002. *(Blank subjects count as 1)*	**Answer:** 4
	Compare	Which month had more total email messages **January, 2002** or **February, 2002**?	**Answer:** February
	Correlate	Did people not send emails on the weekends during the month of **March, 2002**	**Answer:** NO
	Rank	Which subject was most discussed during the month of March	**Answer:** "Historical Record Annual Reviews"
Task Set 2:	**Cluster**	What subjects were discussed during the month of **February, 2002** and had the same total # of email messages?	**Answer:** 1) "Looking For Work" 2) "Use of the AFCA Email Lists"
	Identify	Identify the total number of distinct subjects during the month of December, 2001. *(Blank subjects count as 1)*	**Answer:** 15
	Compare	Which month had more total email messages **December, 2001** or **March 2002**?	**Answer:** December
	Correlate	Did people not send emails on the weekends during the month of **January, 2002?**	**Answer:** NO
	Rank	Which subject was most discussed during the month of **February, 2002**?	**Answer:** "ATCALS to OSS Proposed Implementation Plan"

SESSION C	Strengths/Weakness Tool vs. Text-based interface *(group by Sender or by Subject)*		
Task Set 1:	**Cluster**	Has at least one person from the units, "**49 CS/SCMR**", and "**83 CS/SCYA (AFETS)",** shown a strength by responding to the subject "**ALOD Automatic Lock Out Device**"?	**Answer:** YES
	Identify	Identify maximum of one (1) subject that had a person from unit "**99CS/SCMYM**" respond to it (show a strength).	**Answer:** (1) "Historical Ann..." **OR** (2)"Secure Equip…" **OR** (3) "T.O. Reference"
	Compare	Which unit replied more to "**Anyone Using CAMS GUI?**", "**4 AF/SCM**" or "**31 CS/CMYQ**"?	**Answer:** 4 AF/SCM
	Correlate	Do persons from unit "**31CS/SCMW**" ask more questions (weakness) or reply with answers (strength).	**Answer:** Answers (strength)
	Rank	Which Unit has displayed a strength by replying most to the subject "**crosstalk/bleedover??"?**	**Answer:** 81 TRSS/TSQR
Task Set 2:	**Cluster**	Has at least one person from the units, "**81 TRSS/TSQR**", and "**83 CS/SCYA (AFETS)",** shown a strength by responding to the subject "**CAMS GUI Update**"?	**Answer:** NO
	Identify	Identify maximum of one (1) subject that had a person from unit "**437 CS/SCMR**" respond to (show a strength).	**Answer:** "ATCALS to OSS Proposed Implementation Plan"
	Compare	Which unit replied more to "**ETVS Test Equipment" "83 CS/SCYA (AFETS),** or "**31CS/SCMFG"**?	**Answer:** 83 CS/SCYA (AFETS)
	Correlate	Do persons from unit "**452 SPTG/SCM**" ask more questions (weakness) or reply with answers (strength).	**Answer:** Answers (strength)
	Rank	Which Unit has displayed a strength by replying most on the subject "**LMQCC for ETVS**"?	**Answer:** AFCA/WFLM

SESSION D	Content/Behavior Tool vs. Text-based interface *(group by Sender, Subject or "Received")*		
Task Set 1:	**Cluster**	What subjects were discussed by **SrA's** during the month of **April, 2002**.	**Answer:** 1) "SRD for the UXC-10 / TS-21 Blackjack" 2) "LWC for GRA-4 & PSC-5"
	Identify	Identify the total number of unique subjects for the month of **March, 2002** for those with the rank of **TSgt**.	**Answer:** 3
	Compare	For the month of **April 2002**, did **MSgt's** participate in more conversion on subjects "**Definition of ATCALS**" or "**TVSP**"	**Answer:** Definition of ATCALS
	Correlate	For the month of February, does it appear that **SSgt's** replied more to **"ATCALS to OSS Implementation Plan**" than **SrA's**,	**Answer:** YES
	Rank	For **TSgt's**, what subject had the most discussion in **April, 2002**?	**Answer:** "ALOD Automatic Lock Out Device"
Task Set 2:	**Cluster**	What subjects were discussed by **TSgt's** during the month of **January, 2002**.	**Answer:** AFMQCC 100-3
	Identify	Identify the total number of unique subjects for the month of March for those with the rank of "**MSgt**".	**Answer:** 2
	Compare	For the month of **April, 2002** did **SrA's** participate in more conversion on subjects "**LWC for GRA-4 & PSC-5**" or '**SRD for the UXC-10 / TS-21 Blackjack**"?	**Answer:** "LWC for GRA-4 & PSC-5"
	Correlate	For the month of **February, 2002**, does it appear that **Contr's** (Contractors) replied more to **"AF CEMI for Air Traffic Light Gun**" than **Civ's** (Civilians).	**Answer:** NO
	Rank	For **SMSgt's**, what subject had the most discussion for the month of **February, 2002**?	**Answer:** "ATCALS to OSS Proposed Implementation Plan"

Appendix C: End-User Questionnaire

End-User Questionnaire

A) What is your rank_____

B) Are you currently a member of the career-related listserv under study?
(Ground Radio Systems)

☐ Yes

☐ No

C) If answered "Yes" for question B, how many years have you been a member of the career-related listserv under study? _____

D) Do you currently use Microsoft Outlook on the job to manage work emails?

☐ Yes

☐ No

E) If answered "Yes" for question D, how many years have you been using Microsoft Outlook? _____

ORGDISCOVERY SYSTEM

1) For the tasks performed, the OrgDiscovery system is user friendly.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

2) For the tasks performed, the OrgDiscovery system is easy to use.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

3) Using OrgDiscovery would help me accomplish such tasks more quickly.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

4) Using OrgDiscovery would improve my job performance, given these tasks performed consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

5) Using OrgDiscovery, for the tasks performed, would increase my productivity.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

6) Using OrgDiscovery would improve my effectiveness on the job, given these tasks consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

7) Using OrgDiscovery would make it easier to do my job, given these tasks consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

8) I would find OrgDiscovery useful in my job, given these tasks consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

Please add any additional comments below about the OrgDiscovery System:

_____
_____
_____
_____
_____
_____
_____

MICROSOFT OUTLOOK

1) For the tasks performed, Microsoft Outlook is user friendly.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

2) For the tasks performed, Microsoft Outlook is easy to use.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

3) Using Microsoft Outlook would help me accomplish such tasks more quickly.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

4) Using Microsoft Outlook would improve my job performance, given these tasks performed consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

5) Using Microsoft Outlook, for the tasks performed, would increase my productivity.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

6) Using Microsoft Outlook would improve my effectiveness on the job, given these tasks consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

7) Using Microsoft Outlook would make it easier to do my job, given these tasks consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

8) I would find Microsoft Outlook useful in my job, given these tasks consisted in my job requirements.

Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
☐	☐	☐	☐	☐	☐	☐

9) Which software program do you prefer for the tasks performed.

☐ **Microsoft Outlook**

☐ **OrgDiscovery**

Please add any additional comments below about the Microsoft Outlook:

_____
_____
_____
_____
_____
_____
_____

Bibliography

Ackerman, M.S., (1998). Augmenting Organizational Memory: A Field Study of Answer Garden, *ACM Transactions on Information Systems*, 16(3), 203-224.

----- (1992). AnswerGarden and the Organization of Expertise. *Working Paper CCS TR#127*, MIT Center for Coordination Science.

----- & Malone, T.W. AnswerGarden: A Tool for Growing Organizational Memory. Proceedings of the ACM Conference on Office Information Systems, 31-39.

Adams, D. A., Nelson, R. R., Todd, P.A. Perceived Usefulness, Ease of Use, and Usage of Information Technology: A Replication. *MIS Quarterly*, 16(2), 227-247.

Ahuja, M. K. (1998) Network Structure in Virtual Organizations. *Journal of Computer-Mediated Communication*, 3(4). WWW.ascusc.org/jcmc/vol3/issue4/ahuja.html.

Argote L. (1999). *Organizational Learning*: *Creating, Retaining and Transferring Knowledge*. Norwell, MA: Kluwer Academic Publishers.

Argyris, C. & Schon D.A., (1978). *Organizational Learning: a Theory of Action Perspective*, Reading, MA: Addison-Wesley.

Brown, J.S & and Duguid, P. "Organizational Learning and Communities of Practice: Toward a Unified View of Working, Learning, and Innovation (1995). *Organizational Science*, 2(1), 40-57.

Bruha, I. (2000). From Machine Learning to Knowledge Discovery: Survey of Preprocessing and Postprocessing. *Intelligent Data Analysis*, 4(3/4), 363-374.

Chait, L. P. (2000), Creating a Successful Knowledge Management System. *IEEE Engineering Management Review*. 92-95.

Chen, H., Titkova, O., Orwig, R., & Nunamaker, J. F. (1998), "Information Visualization for Collaborative Computing," *IEEE Computer*, 31(8), 75-82.

Collins, B.M. (1993) Data visualization — has it all been seen before? *In: Animation and Scientific Visualization Tools and Applications.* London: Academic Press, 3–28.

Conklin, J. & M.L. Begelman, (1988), gIBIS: A Hypertext Tool for Exploratory Policy and Discussion. Proceedings of the Conference on Computer Supported Cooperative Work, 8, (proceedings not pages).

Davenport, T.H., & Prusak L. (1998). *Working Knowledge: How Organizations Manage What They Know*. Boston, MA: Harvard Business School Press.

Davis, F. D. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 318-340.

-----, Bagozzi, R. P., & Warshaw, P. R. (1989) "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science*, 35, 982-1003.

DeTienne, K. B., & Jackson, L. A. (2001). Knowledge management: understanding theory and developing strategy. *Competitiveness Review*, 1, 1-11.

Doll, W. J., Torkzadeh, G. (1998). The Measurement of End-User Computing Satisfaction. *MIS Quarterly*, 12(2) 258-274.

Donath, J., Karahalios, K., & Viégas, F. (1999), Visualizing Conversation, *Journal of Computer-Mediated Communication*, 4(4), http://www.ascusc.org/jcmc/vol4/issue4/donath.html

Drucker, P.F. (1998). The Coming of a New Organization. In *Harvard Business Review on Knowledge Management* (pp.1-20). Boston MA: Harvard Business School Publishing.

Dumas, J. S., & Redish, J.J. (1993). *A Practical Guide to Usability Testing*. Norwood NJ: Ablex Publishing Corporation.

Elofson, G., & Konsynski, B., (1991) Delegation Technologies: Environmental Scanning with Intelligent Agents. *Journal of Management Information Systems*, 8(1), 37-62.

Finch, S., & Mikheev, A. (1997). A Workbench for Finding Structure in Texts. Paper presented at *Applied Natural Language Processing,* Washington D.C.

Gore, C., & Gore, E. (1999). Knowledge management: the way forward. *Total Quality Management*, 4/5, 554-561.

Goodman, P.S & Darr, E.D. (1998). Computer-Aided Systems and Communities: Mechanisms for Organizational Learning in Distributed Environments, *MIS Quarterly*, 22(4), 417-441.

Grayson, M.A., Biros, D.P, Ward, M., and Chen, H., "Air Force Organizational Memory Information System: A Computer Mediated Communications Perspective," Proceedings of the IASTED International Conference on Information and Knowledge Sharing, Nov 2002, 54-60.

Hackbarth, G., & Grover, V. (1999). The Knowledge Repository: Organizational Memory Information Systems. *Information Management Systems*, 16(3), 21-30.

Hahn, J. & Subramani, M. (2000). "A framework of knowledge management systems: issues and challenges for theory and practice," *Proceedings of the International Conference on Information Systems*, 21.

Huhns, M.N, Mukhopadhyay, U., Stevens, L.M., & Bonnel, R.D. (1989). DAI for Document Retrieval: The Minds Project, in Distributed Artificial Intelligence, L. Gasser and M.N Huhns (Eds.), 2, San Mateo, CA: Morgan Kaufman, 249-283.

Jeffries, R., Miller, J. R., Warton, C., Uyeda, K. (1991) User Interface Evaluation in the Real World: A Comparison of Four Techniques, *Human factors in computing systems conference proceedings on Reaching through technology*, p.119-124.

Jordan, P. K. (1998). *An Introduction to Usability.* Bristol, PA: Taylor & Francis Inc.

Klein, M. (1993). Capturing Design Rationale in Concurrent Engineering Teams, *Computer*, 26(1), 39-47

Kohonen, T. (1995) Self-organized Maps, chapter 3, Spinger-Verlag, Berlin Heidelberg.

Konda, S. Monarch, I., Sargent P., & Subrahmanian E., "Shared Memory in Design," Report EDRC 05-56-91, Engineering Design Research Center Carnegie Mellon University, Pittsburgh, PA, 1991.

Levitt, B. & March, J.G. (1988). Organizational Learning, Annual Review of Sociology, 14, 319-340.

Lim, K. K., Ahmed, P.K., & Zairi, M. (1999). Managing for quality through knowledge management. *Total Quality Management*, 4/5, 615-622.

Lubit, R. (2001) Tacit knowledge and knowledge management: the keys to sustainable competitive advantage. *Organizational Dynamics*, 3, 164-179.

Lynch, K. J., & Chen, J. (1992). Knowledge Discovery from Historical Data: An Algorithmic Approach. *Proceedings from the Hawaii International Conference on System Sciences*, 25(3), 70-79.

Macintosh, A., & Filby, I., Kingston, J. (1999). Knowledge management techniques: teaching and dissemination concepts. *International Journal of Human-Computer Studies*, 3, 549-566.

Markus, M. L. (2001). Toward a Theory of Knowledge Reuse: Type of Knowledge

Reuse Situations and Factors in Reuse Success. *Journal of Management Information Systems*, 18(1), p.57

Mascarenhas, B. (1989). Transitional Linkages and Strategy," in International Strategic Management, A.R Negandhi and J. Liebowitz (Eds.), Lexington, MA: Lexington Books, 53-67.

McCormick, B.H, DeFanti, T.A., & Brown, M.D. (1987). Visualization in scientific computing — a synopsis, *Computer Graphics* 21(6), 1–10.

Morrison, J. (1993). Team Memory: Information Support for Business Teams. *Proceedings of the Hawaii International Conference on System Science*, 26(4), 122-131.

Morse, E., & Lewis, M. (2000). Evaluating visualizations: using a taxonomic guide. *International Journal on Human-Computer Studies*, 53, 637-662.

Nasukawa, T.; & Nagano, T. (2001). Text analysis and knowledge mining system. *IBM Systems Journal*, 40(4), p967-984.

Nelson, K. M., Cooprider, J. G. (1996). The contribution of shared knowledge to IS group performance. *MIS Quarterly*, 4, 409-433.

Nonaka, I. "A Dynamic Theory of Organizational Knowledge Creation" (1994) *Organizational Science*, (5)1, 14-37.

----- & Konno, N. (1998). "The concept of 'Ba': Building a foundation for knowledge creation," *California Management Review*. 40(3), 40-54.

----- (1998). The Knowledge-Creating Company. In *Harvard Business Review on Knowledge Management* (pp.21-45). Boston MA: Harvard Business School Publishing.

Nunamaker, J. F., Dennis, A. R., Valacich, J. S., Vogel, D. R, & George, J. F. (1991). "Electronic meeting systems to support group work," *Communications of the ACM*. 34(7), 40-61.

Rao, R. & Sprague, R.H. (1998). Natural Technologies for Knowledge Work: Information Visualization and Knowledge Extraction. *Journal of Knowledge Management*, 2, 2, 70-80.

Reddy, Y. (1993). Computer Support for Concurrent Engineering. *Computer*, 26(1), 12-16.

Reich, B.H., & Benbasat, I. (2000). Factors that influence the social dimension of

alignment between business and information technology objectives. *MIS Quarterly*. *24*(1), 81-112.

Sandoe, K., & Olfman, L. (2001). Anticipating the Mnemonic Shift: Organizational Remembering and Forgetting in 2001. *Proceedings of the International Conference on Information Systems,* 21, 127-137.

Sekine S., & Grishman, R. (1995). A Corpus-based Probabilistic Grammar with Only Two Non-terminals. International Workshop on Parsing Technology, 4,

Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Reading MD: Addison Wesley Longman, Inc.

Sproull, L. & Kiesler S. (1991). Computers, network, and work. Scientific American, 265(3), 116-127.

Stein, Eric W., & Zwass, Vladimir. (1995). Actualizing Organizational Memory with Information Systems, *Information Systems Research*, *6*(2), 85-118.

Subramani, M. & Hahn, J. (2000). Examining the effectiveness of electronic group communication technologies: The role of the conversation interface," presented at *Academy of Management Conference*. http://www.jungpil.com.

Swap, W., Leonard, D.; Shields, M., Abrams, L. (2001) Using mentoring and storytelling to transfer knowledge in the workplace. *Journal of Management Information Systems,* 18, 95-115.

Szajna B. (1994). Software Evaluation and Choice: Predictive Validation of the Technology Acceptance Model. MIS Quarterly, 18(3) 319-325.

Tolle, K. M., & Chen, H. (2000). Comparing Noun Phrasing Techniques for Use with Medical Digital Libraries. *Journal of the American Society for Information Science,* Special Issue on Digital Libraries, 51(4), 352-370.

Trauth, E.M., & Jessup, L.M. (2000). Understanding Computer-Mediated Discussions: Positivists and Interpretative Analyses of Group Support System Use, *MIS Quarterly*, 9(1), 43-79.

United State Air Force. (1999) Air Force Instruction 33-119, Section 5.3

Walsh, J.P., & Ungson, G.R. (1991) Organizational Memory, *Academy of Management Review*, 16(1), 57-91.

Wehrend, S., & Lewis, C. (1990).A problem-oriented classification of visualization techniques. *Proceedings of the IEEE Visualization*, pp. 139-143.

Xiong, R. & Donath, J. (1999). PeopleGarden: Creating Data Portraits for Users. *Proceedings Annual ACM Symposium on User Interface Software and Technology*, 12, 37-44.

Zhu B., & Chen H. (2001). Social Visualization for Computer-Mediated Communication: A Knowledge Management Perspective, *Proceedings of the Workshop On Information Technologies And Systems*, 11, 23-28.

Vita

2d Lt Marc Grayson graduated from the University of Arizona with a Bachelor of Science in Comp uter Science in May, 2001. He was commissioned through Air Force ROTC Detachment 020. He is an alumnus of the Artificial Intelligence (AI) lab in the Department of Management Information Systems, College of Business and Public Administration, at the University of Arizona. His first assignment was to the Air Force Institute of Technology (AFIT), Wright-Patterson AFB, OH from 2001 to 2003. His Second assignment was to the Air Force Communications Agency (AFCA) at Scott AFB, IL, from 2003.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
25-03-2003	**Master's Thesis**	Aug 2001 – Mar 2003

4. TITLE AND SUBTITLE

VISUALIZATION OF CAREER-RELATED COMPUTER-MEDIATED COMMUNICATION FOR INCREASED KNOWLEDGE MANAGEMENT

5a. CONTRACT NUMBER
N/A

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)

Grayson, Marc A., 2d Lt., USAF

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
N/A

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 P Street, Building 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GIR/ENV/03-05

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFCA/WFLM
Attn: Ms. Susan Woodson
203 W Losey St, Rm# 2100 DSN: 779-6484
Scott AFB, IL 62225-5222 e-mail: Susan.Woodson@scott.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)
N/A

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
N/A

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Retention and retrieval of organizational memory has been the concentration of many conceptualized models of an organizational memory information system (OMIS). This thesis presents an extended view for system development of an OMIS from a knowledge management perspective. The United States Air Force maintains various career-related mailing lists (listservs) for information technology (IT) specialists sponsored by the Air Force Communications Agency (AFCA). AFCA has realized the importance of monitoring the communication for patterns in content and behavior. This thesis details an experimental study, which includes a repository of computer-mediated communication (CMC) of IT specialists, analyzed by software created for this study, the OrgDiscovery system. This system is designed to visualize the content and behavior patterns of computer-mediated communication. The purpose of this study is to show that visualization of mailing list communication provides a more usable method to make conclusions about the participants of mailing lists versus the text-based Microsoft Outlook. M.S. Outlook is the mail program currently being used by management to store and review mailing list emails.

15. SUBJECT TERMS

Knowledge Visualization, Computer-Mediated Communication, Organizational Memory Information Systems, Listserv

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Mark Ward, Maj, USAF (ENV)
U	U	U	UU	109	19b. TELEPHONE NUMBER *(Include area code)* (937) 785-3636 x4742; e-mail: mark.ward@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18